A POLYNUCLEOTIDE FUNCTIONALLY CODING FOR THE LHP PROTEIN FROM MYCOBACTERIUM TUBERCULOSIS, ITS BIOLOGICALLY ACTIVE DERIVATIVE FRAGMENTS, AS WELL AS METHODS USING THE SAME

INS C1 5 1

BACKGROUND OF THE INVENTION

Field of the Invention

The present invention is directed to a polynucleotide comprising an open reading frame coding for a polypeptide from Mycobacterium tuberculosis, named LHP (also referred as CFP-10) capable of inducing an immune response in a host, said LHP is placed under the control of its own regulation signals which are functional in mycobacteria, specially in mycobacteria belonging to the Mycobacterium tuberculosis complex and also in fast growing mycobacteria such as Mycobacterium smegmatis and also in E. coli. The Mycobacterium tuberculosis complex has its usual meaning, i.e. the complex of mycobacteria causing tuberculosis which are Mycobacterium tuberculosis, Mycobacterium bovis, Mycobacterium africanum, Mycobacterium microti and the vaccine strain M. bovis BCG.

The invention is also directed to the polypeptide LHP encoded by lhp and most preferably to suitable antigenic portions of LHP as well as to oligomeric polypeptides containing more than one unit of LHP or an antigenic portion of LHP. The invention concerns also immunogenic and vaccine compositions containing a polypeptide or an oligomeric polypeptide such as defined above or live recombinant attenuated mycobacteria transformed with a polynucleotide according to the present invention. The invention also concerns antibodies directed specifically against such polypeptides that are useful as diagnostic reagents. In another embodiment, the present invention is directed to a polynucleotide carrying the natural regulation signals of lhp which is useful in order to express heterologous proteins in

mycobacteria as well as functionally active regulatory polynucleotides derived from said regulatory region. Finally, the present invention is directed to oligonucleotides comprising at least 12 consecutive nucleotides which are useful as reagents for detecting the presence of Mycobacterium tuberculosis in a biological sample.

Related Prior Art

Mycobacterium tuberculosis and M. bovis cause tuberculosis, a disease which currently kills three million people each year. The virulence of pathogenic mycobacteria is associated with their ability to parasitize and survive within phagocytic cells. Little is known about mechanisms governing gene expression during the intracellular growth stage. This issue is of prime importance as the intracellular stage of pathogenic mycobacteria can be viewed as an adaptative process, involving transcriptional regulatory mechanisms. Mycobacterial genes affecting intracellular growth and virulence are being actively sought (Collins, 1996; Collins, 1995, Quinn, 1996). Using subtractive genomic hybridization between virulent M. bovis and the attenuated vaccine strain M. bovis BCG, Maheiras et al. (Maheiras et al., 1996) identified three regions of difference (RD1 to RD3). RD1 was detected in all strains of M. tuberculosis and M. bovis tested but is absent in all BCG substrains, suggesting that it may be an important determinant of virulence.

The orf1C gene, encoding the early secreted antigenic target 6kDa (ESAT-6) lies within RD1. The ESAT-6 protein is a major T-cell antigen which has been purified from M. tuberculosis short-term culture filtrates (Harboe et al., 1996; Sorensen et al., 1995). Purified ESAT-6 stimulates the production of gamma interferon from mice memory immune T lymphocytes and may contribute to the development of antituberculous immunity (Andersen et al., 1995; and U.S. Patent Application filed on June 5, 1995).

The Mycobacterium genus encompasses more than 70 recognized bacterial species including M. tuberculosis and M. leprae, the agents of tuberculosis and leprosy respectively. The development of effective prophylactic vaccine and specific

diagnostic reagents is a priority to control the extension of mycobacterial infections. In that context, mycobacterial protein antigens are extensively screened upon their ability to induce B- and T-cell reactivity. Obtention of purified proteins from slow growing pathogenic mycobacteria is labor-intensive and requires important containment facilities. Alternatively, many immunological studies of mycobacterial antigens have been conducted with E. coli-expressed recombinant molecules. However, problems related to lipopolysaccharide (LPS) contamination are frequently encountered. Moreover, post-translational modifications such as proteolytic processing, intern removal, lipid acylation and glycosylation of proteins have been reported to occur in mycobacteria. Such modification cannot be mimicked in E. coli and may influence dramatically the stability, antigenicity and the immunogenicity of the peptide chain. Thus, it was recently postulated that site-specific mannosylation protects the M. tuberculosis 19kDa lipoprotein antigen against proteolysis (Hermann). Accordingly, there is a great need in the art of suitable protein expression systems allowing the preparation of mycobacterial immunogenic polypeptides that are useful for diagnostic and vaccine purposes.

Summary of the Invention

Now, the inventors have discovered a polynucleotide carrying the regulatory expression signals of the ESAT-6 protein as well as an open reading frame coding for a new antigenic protein from Mycobacterium tuberculosis that they have named LHP.

The LHP polypeptide of the invention share a great similarity with a Mycobacterium tuberculosis peptide described in the PCT Application No. WO 97/09429 or in the PCT Application No. WO 97/09428 (Corixa Corporation) a partial sequence of which is disclosed in those patent applications.

The present inventors have characterized the portions of the polynucleotide according to the invention that are functional in mycobacteria in order to allow the expression of LHP, as well as the expression of an heterologous polypeptide that is placed under

the control of said regulatory region contained in the polynucleotide according to the present invention.

More specifically, the inventors have located the transcription initiation sites of the lhp/orflC operon using M. tuberculosis RNA and have precisely mapped the portions of the regulatory region of the lhp/or lC operon that are functional in bacteria in general, being functionally active in E. coli as well as in mycobacteria. Further, the inventors have mapped the portions of the polynucleotide according to the present invention that are functionally active in slow growing mycobacteria, such as bacteria belonging to the Mycobacterium tuberculosis complex, and in fast-growing mycobacteria, such as M. smegmatis.

Further, the present inventors have used the functionally active portions of the regulatory region of the lhp/orflC operon for expressing a polypeptide heterologous with respect to said regulatory region.

In a specific embodiment, the present inventors have constructed a mycobacterial expression vector allowing production of recombinant proteins tagged by a stretch of six histidine. Such vector enables production of virtually any polypeptide in a mycobacterial context and allows easy purification of native proteins by Immobilized metal affinity chromatography. Additionally, the availability of monoclonal antibody directed against the (His)6 polypeptide facilitates the detection of proteins for which no specific immune reagent are available. This system is very useful for biochemical and immunological characterization of mycobacterial proteins.

Accordingly, given its high level and constitutive expression of the regulatory polynucleotide according to the present invention in mycobacteria, said promoter is used to construct a novel mycobacterial expression/purification system.

This vector designated pIPX30, allows versatile gene fusions to produce histidine-tagged proteins in mycobacteria. Additionally, the high affinity of polyhistidine for immobilized metal ions enables one-step chromatographic isolation of native, histidine-tagged polypeptides. As a validation of the system, the inventors have performed the expression of recombinant DES(Histidine)6 M. tuberculosis protein antigen and its immunodetection from M. smegmatis cultures.

Thus, the present invention is directed to a polynucleotide comprising a functional portion of the regulatory region of the lhp operon and to its use in a recombinant expression vector carrying a polynucleotide encoding a polypeptide of interest.

The invention also concerns recombinant expression vectors containing a polynucleotide according to the invention, and more specifically a polynucleotide carrying one of the regulatory polynucleotides characterized by the inventors.

The invention is also directed to recombinant cell host containing a polynucleotide or a recombinant vector as defined above.

In another embodiment, are also part of the present invention the entire LHP antigenic polypeptide as well as particular antigenic portions of the LHP polypeptide that have been identified by the inventors.

A further embodiment of the present invention is directed to oligomeric polypeptides that contain at least one unit of an antigenic portion of the LHP polypeptide, that are useful as immunogenic molecules. Consequently, the present invention concerns also immunogenic compositions as well as vaccine compositions that are useful to diagnose and to prevent an infection by mycobacteria belonging to the M. tuberculosis complex, and more specifically by Mycobacterium tuberculosis in humans and animals.

Another object of the present invention consists in a polyclonal or a monoclonal antibody directed specifically against the LHP polypeptide or an antigenic portion thereof.

The present invention concerns also methods and corresponding kits containing either a polynucleotide, polypeptide or an antibody according to the invention in order to perform the diagnosis of an infection with Mycobacterium tuberculosis in a biological sample.

The invention pertains to immunogenic and vaccine compositions containing at least a polypeptide or a recombinant cell host expressing the LHP polypeptide, preferably in combination with the ESAT-6 antigenic protein and also to vaccine compositions containing live non pathogenic recombinant cell hosts expressing these polypeptides.

Finally, the invention also describes a M. tuberculosis complex strain, deleted for the lhp / esat-6 region, which is useful for testing the importance of both antigenes lhp and esat-6 in the protection or development of the disease after infection. Such a strain referred as H37Rv Δ (lhp esat-6). This strain derives from the Mycobacterium tuberculosis strain of reference H37Rv, which is virulent in human. H37Rv Δ (lhp esat-6) presents a complete deletion of the lhp gene coding for the LHP protein (also called CFP-10 or ORFX) and a partial deletion of the esat-6 gene (figure 13). ESAT-6 and LHP are antigenes secreted by Mycobacterium tuberculosis. H37Rv Δ (lhp esat-6) was constructed by replacing (homologous recombination) the PstI fragment of 1.1 kpb containing lhp and esat-6 by a pstI fragment of 1.3 kbp containing kanamycine resistant gene. H37Rv Δ (lhp esat-6) has been deposited on June 29, 1998 at the CNCM (Institut Pasteur, 25 rue du Docteur Roux, 75724 Paris cedex 15, France) under the accession number I-2047.

Brief Description of the Figures

Figure 1 - E. coli strain [pIPX26] : Functional and structural features contained in the Kpn Bam HI insert.

Main features of the nucleotide insert contained in plasmid pIPX26 that has been deposited at the CNCM on May 14, 1996 under the Accession Number I-1706. This insert contains the whole polynucleotide carrying the lhp-or IC operon. pIPX26 is a shuttle cloning vector (E. Coli-mycobacteria) of the pPV24 kind conferring kanamycin resistance and carrying a DNA insert at the unique cloning sites KpnI (Asp718) and BamHI). This DNA insert is a 1282 bp DNA fragment form Mycobacterium tuberculosis H37Rv, which has been generated by PCR amplification using the following pair of primers:

ESB-1 (5'-GGGGGGATCCGGTACCAGGTGACGTCGTTGTTCAGCCAG-3') (SEQ ID NO:17)

and

ESB-2 (5'-GGGGGGTACCGGATCCTCGTAGTCGGCCGCCATGACAAC-3') (SEQ ID NO:18)

and by digestion with the restriction enzymes Asp718 and BamHI. This DNA fragment carries the open reading frame referenced orfX (which is the lhp gene), the ESAT-6 (also referred to as or lC) gene and its own transcription terminator. This DNA fragment comprises also a promoter activity and transcription start sites allowing gene expression, including lhp and ESAT-6 (orflC) in M. smegmatis and M. bovis-BCG. α and β represent respectively the transcrition start site in Mycobacterium tuberculosis Mt 103 and Mycobacterium smegmatis mc2 155.

When plasmid pIPX26 is transferred in M. smegmatisI and M. bovis-BCG, the ESAT-6 protein, which is normally absent from these mycobacterial strains is expressed. This ESAT-6 heterologous expression is detected by Western blot with the monoclonal antibody Hyb 76-8 on protein extracts.

Figure 2 - Gene arrangement upstream from the M. tuberculosis orflC gene and lacZ gene fusions used in this study.

The 1.1kb PstI fragment from pAA249 was blunted with T4 DNA polymerase in the conditions described by the supplier (New England Biolabs, MA USA). Insert of this DNA fragment into T4-blunted, SnaBI-digested pJEM13 and pJEM14 resulted in pIPX15 and pIPX16 respectively. Oligonucleotide pairs
OF1 (5'-GGGGGGATCCCAGGTGACGTCGT TGTTCAGC-3') and
OB1 (5'-GGGGGGTACCACGGTGACGTCGTTGTTCAGC-3'), OF1 and
OB2 (5'-GGGGGGTACCAACGGTGACGTCGTTGTTCAGC-3') together with
PE-1 (5'-GGGGGGTACCGGGTGGCCGGGAAGTCTGTTG-3') and
PE-4 (5'-GGGGGGATCCCTGCAGCAGGTGACGTCGTTG-3') and E64 (5'-CCCTGCAACGAACCTGCCGTCGACTCCACC-3') were used for PCR amplification from pIPX61. Plasmids pIPX45, pIPX46 and pIPX18 were obtained by insertion of BamHI/Asp718-digested PCR fragments into the corresponding sites in pJEM13 and pJEM15. Stem/loops represent probable transcription terminators and open triangles indicate 18bp tandem repeats upstream from lhp. β-galactosidase activities in M. smegmatis are 85 U ± 21 for pIPX15, 1789 U ± 75 for pIPX16, 77

U ± 6 for pIPX46, 1010 U ± 29 for pIPX47, and 36 U ± 4 for pIPX18. Results of β-galactosidase assays and means and standard deviations of three measurements and were determined in M. smegmatis as described in (Timm et al., 1994).

Figure 3 - Main features of the nucleotide insert contained in plasmid pIPX61.

pIPX61 has been deposited at the CNCM on May 14, 1996 under the Accession Number I-1705. The p Bluescript 11 KS (+/-) vector (Invitrogen) has been used as a basis for constructing pIPX61. α and β represent respectively the transcrition start site in Mycobacterium tuberculosis Mt 103 and Mycobacterium smegmatis mc2 155.

Figure 4 - Nucleotide and amino acids sequence features upstream from the M. tuberculosis orf1C start codon.

(nucleotides 1-867 of SEQ ID NO:1) (SEQ ID NO:5)

(A) nucleotide sequence and deduced amino acid sequence of lhp. A potential ribosome binding site (RBS) upstream from the predicted start codon is underlined. Transcriptional start sites identified in M. tuberculosis (+1 Mtb) and in M. smegmatis (+1 Ms) are indicated by triangles.

(B) peptide sequence similarity between the predicted M. tuberculosis lhp gene (SEQ ID NO:28) product and the M. leprae L45 seroreactive protein antigen (Accession Number X90946).

Figure 5 – Mapping of the lhp-orf1C promoter activity.

(A) Primer extension mapping of the transcriptional start sites (T1, T2 and T3) in M. tuberculosis. Reverse transcription was performed as described in (Berthet et al., 1995) using the E64 oligonucleotide (SEQ ID NO:24)
(5'-CCCTGCAACGAACCTGCCGTCGACTCCACC-3') with (lane 1) or without (lane 2) RNA. The DNA ladder was generated by sequencing pIPX61 with E64 using the T7 sequencing kit (Pharmacia Biotech).

nucleotides 392-482 of SEQ ID NO: 1

(B) Structural features of the M. tuberculosis orf1C promoter.

(C) Primer extension mapping of the transcriptional start sites (S1 and S2) in M. smegmatis transformed with pIPX16. Experimental conditions were the same as described in (A).

Figure 6 - Analysis of the *lhp/orf1C* messenger RNA transcript.

Total RNA was extracted from *M. tuberculosis* broth cultures on day 5 (lane 1 and 2), day 9 (lane 3), day 13 (lane 4) and day 16 (lane 5). Total RNA (5µg) was separated on 1% agarose gel supplemented with formamide/formaldehyde and processed for Northern blotting as described in (Sambrook et al., 1989). Hybridization was carried out using the radiolabeled ESA-A probe (See Figure 2). Autoradiography was performed for 4 (lane 1) to 24 hours (lane 2 to 5).

Figure 7 - Features of the pIPX30 expression/tagging plasmid.

Plasmid pIPX30 is derived from plasmid pPV24 and is a shuttle plasmid possessing the following features:

(1) the origin of replication of pAL5000 for propagation in mycobacteria, the origin of replication from vector pUC19 allowing its propagation in E. coli, the aph selection gene conferring resistance to kanamycin;

(2) the promoter region of lhp and ESAT-6 from M. tuberculosis, functionally active in slow growing (M. Tuberculosis, M. bovis-BCG, etc.) and in fast growing mycobacteria (M. Smegmatis); (SEQ ID NOS: 29 and 30)

(3) an expression cassette consisting in: Shine-Dalgarno site/ATG from plasmid pJEM15, three cloning sites (BamHI, KpnI, PstI), a DNA fragment coding for six Histidine, two translation stop codons and the transcription terminator form ESAT-6.

Plasmid pIPX30 has been constructed by digestion of plasmid pPV24 with KpnI/PstI, then treated by phage T4 DNA polymerase and then by insertion of an expression cassette having blunt-ended at 5' and 3' ends.

Plasmid pIPX30 allows the production of proteins having a six Histidine stretch on their NH2 extremity. This feature facilitates their purification by affinity chromatography on columns endowed with immobilized metal ions (IMAC).

Figure 8 - Beta-galactosidase activities of M. smegmatis clones containing pIPX34 or positive (pJN30) and negative (pJEM13) control vectors.

Figure 9 - Immunodetection of DES-(His)6 in Mycobacterium smegmatis protein extracts.

Lanes 1-2: revelation with an anti-DES polyclonal antiserum. Lanes 3-4: revelation with a monoclonal antibody directed against X(His)6.

Lanes 1 and 3: mc2 155 w+ (wild type)

Lanes 2 and 4: mc2 155 [pIPX30-DES]

Figure 10 - Map of plasmid pPV24.

PPV24 is a shuttle plasmid (E. Coli - mycobacteria). This plasmid has been constructed in two steps:

(a) a large portion of the ampiciline resistance coding gene as well as the neighboring non-useful sequences of plasmid pUC18 (NdeI + BsaI fragment) have been replaced by the kanamycin resistance gene from pUC4K (PstI fragment) which also express in mycobacteria. The resulting vector is pPV8 (2.8kb);

(b) (b) the minimal origin of replication of the mycobacterial plasmid pAL500 (EcoRV + HpaI fragment) has been cloned at the StuI site from pPV8. The final vector is pPV24 (5.4kb), which carries the multiple cloning site from pUC18 and allows the direct detection of recombinant host cells on culture medium supplemented with X-Gal.

Figure 11 - pPX1 is a shuttle cloning vector (E. Coli - mycobacteria) of the pPV24 kind, which confers kanamycin resistance and possessing a 855bp insert at the BamHI unique cloning site.

The 855bp insert from Mycobacterium tuberculosis H37 Rv is generated by PCR amplification using the following primer pair:

ESB-1 (SEQ ID NO:17)

(5'-GGGGGGATCCGGTACCAGGTGACGTCGTTGTTCAGCCAG-3')

PO-1

(5'-GGGGGGATCCTCAATGGTGATGGTGATGGTGGAAGCCCATTTGCGAG GACAGCGC-3') (SEQ ID NO: 26)

and then by digestion with the restriction enzyme BamHI. This DNA fragment contains the open reading frame referenced orfX (which is the lhp gene) fused to a DNA stretch coding for six Histidine. This DNA fragment carries a promoter region and transcription start sites, allowing gene expression in Mycobacterium smegmatis and Mycobacterium bovis-BCG and Mycobacterium tuberculosis. α and β represent respectively the transcrition start site in Mycobacterium tuberculosis Mt 103 and Mycobacterium smegmatis mc2 155.

Figure 12- Detection of M. tuberculosis CFP-10.

(A) Protein content of the M. tuberculosis ST-CF analyzed by SDS-PAGE and silver staining (first lane) and corresponding purified low molecular weight fractions (following lanes) analyzed by SDS-PAGE. Fraction number 4 contained LHP which N-terminal sequence is indicated (arrow).

(B) Separation of recombinant ESAT-6 and rLHP under SDS-PAGE conditions.

Figure 13- Allelic exchange at the lhp/esat-6 loci.

H37Rv Δ (lhp esat-6) derives from the Mycobacterium tuberculosis strain of reference H37Rv, in which a complete deletion of the lhp gene and a partial deletion of the esat-6 gene has been realized. H37Rv Δ (lhp esat-6) was constructed by replacing (homologous recombination) the PstI fragment of 1.1 kpb containing lhp (SEQ ID NO 26) and esat-6 by a pstI fragment of 1.3 kbp containing kanamycine resistant gene. The sequence shown is the deleted sequence.

Detailed Description of the Preferred Embodiments

The present inventors have discovered a new polynucleotide and have shown that said polynucleotide contained a whole operon consisting in a regulatory region containing a functional promoter and a functional ribosome binding site that drives the expression of two structural genes respectively encoding a new polypeptide named LHP and an already known polypeptide named ESAT-6.

Further, the inventors have discovered that the two structural genes are co-transcribed under the control of the said promoter region.

The inventors have further characterized the LHP polypeptide as being a polypeptide produced and excreted by Mycobacterium tuberculosis. The inventors have also demonstrated that the polypeptide LHP was produced simultaneously with the antigenic polypeptide ESAT-6 in Mycobacterium tuberculosis. As shown herein by the inventors, via a micro sequencing method of the peptides excreted in the culture medium supernatant of Mycobacterium tuberculosis, the LHP polypeptide is secreted by said pathogenic bacterium.

Moreover, the present inventors have shown that the regulatory region located at the 5' end of the open reading frame coding for LHP can be successfully used to drive the expression of an heterologous polynucleotide as regards to LHP in a recombinant cell host.

For this purpose, the inventors have designed three plasmids containing the regulatory region of lhp and ESAT-6 (orflC), namely plasmids pIPX30, pIPX26 and pPX1.

pIPX30 has been deposited at the CNCM (Collection Nationale de Cultures de Microorganisms) on February 13, 1997, under the accession number I-1845. A map of plasmid pIPX30 is shown on Figure 7.

pIPX26 has been deposited at the CNCM on May 14, 1996, under the Accession Number I-1706. A map of pIPX26 is shown on Figure 1. pIPX26 has been designed using a specific plasmid construct named pPV24 (see Figure 10), that has been deposited at the CNCM on May 14, 1996, under the Accession Number I-1704, which is part of the invention.

pPX1 contains the regulatory region of lhp/orf1C, the open reading frame coding for lhp, the stop codon of which has been replaced by a polynucleotide coding for six histidine and ending with a stop codon. pPX1 has been deposited at the CNCM on May 14, 1996 under the Accession Number I-1707. A map of plasmid pPX1 is shown on Figure 11.

Taking into account that neither the regulatory region sequence nor the LHP encoding nucleotide sequence were found to have a strong homology with already known nucleotide sequences and then taking into account of their uniqueness in mycobacteria, a further object of the present invention consists in polynucleotides derived from the polynucleotide containing the lhp/orf1C operon, or alternatively a polynucleotide hybridizing under stringent hybridization conditions with the polynucleotide containing the lhp/orf1C operon, which are useful as primers or probes in order to detect specifically a bacterium of the Mycobacterium tuberculosis species in a biological sample.

Thus, the present invention is directed to a purified polynucleotide wherein said polynucleotide is chosen from the group consisting of:

(a) a polynucleotide comprising the following nucleotide sequence of SEQ ID NO 1:

CTGCAGCAGGTGACGTCGTTGTTCAGCCAGGTGGGCGGCACCGGCGGCGG
CAACCCAGCCGACGAGGAAGCCGCGCAGATG
GGCCTGCTCGGCACCAGTCCGCTGTCGAACCATCCGCTGGCTGGTGGATC
AGGCCCCAGCGCGGGCGCGGGCCTGCTGCG

CGCGGAGTCGCTACCTGGCGCAGGTGGGTCGTTGACCCGCACGCCGCTGA
TGTCTCAGCTGATCGAAAAGCCGGTTGCCC
CCTCGGTGATGCCGGCGGCTGTTGCCGGATCGTCGGTGACGGGTGGCGCC
GCTCCGGTGGGTCCGGGAGCGATGGGCCAG
GGTTCGCAATCCGGCGGCTCCACCAGCCCGGGTCTGGTCGCGCCGGCACC
GCTCGCGCAGGAGCGTGAAGAAGACGACGA
GGACGACTGGGACGAAGAGGACGACTGGTGAGCTCCCGTAATGACAACA
GACTTCCCGGCCACCCGGGCCGGAAGACTTG
CCAACATTTTGGCGAGGAAGGTAAAGAGAGAAAGTAGTCCAGCATGGCAG
AGATGAAGACCGATGCCGCTACCCTCGGGC
AGGAGGCAGGTAATTTCGAGCGGATCTCCGGCGACCTGAAAACCCAGATC
GACCAGGTGGAGTCGACGGCAGGTTCGTTG
CAGGGCCAGTGGCGCGGCGCGGCGGGGACGGCCGCCCAGGCCGCGGTGG
TGCGCTTCCAAGAAGCAGCCAATAAGCAGAA
GCAGGAACTCGACGAGATCTCGACGAATATTCGTCAGGCCGGCGTCCAAT
ACTCGAGGGCCGACGAGGAGCAGCAGCAGG
CGCTGTCCTCGCAAATGGGCTTCTGACCCGCTAATACGAAAAGAAACGGA
GCAAAAACATGACAGAGCAGCAGTGGAATT
TCGCGGGTATCGAGGCCGCGGCAAGCGCAATCCAGGGAAATGTCACGTCC
ATTCATTCCCTCCTTGACGAGGGGAAGCAG
TCCCTGACCAAGCTCGCAGCGGCCTGGGGCGGTAGCGGTTCGGAGGCGTA
CCAGGGTGTCCAGCAAAAATGGGACGCCAC
GGCTACCGAGCTGAACAACGCGCTGCAGAACCTGGCGCGGACGATCAGCG
AAGCCGGTCAGGCAATGGCTTCGACCGAAG
GCAACGTCACTGGGATGTTCGCATAGGGCAACGCCGAGTTCGCGTAGAAT
AGCGAAACACGGGATCGGGCGAGTTCGACC
TTCCGTCGGTCTCGCCCTTTCTCGTGTTTATACGTTTGAGCGCACTCTGAG
AGGTTGTCATGGCGGCCGACTACGA

(b) a polynucleotide comprising the following nucleotide sequence of SEQ ID NO 2, starting at its 5' end with the nucleotide in position 1 of SEQ ID NO 1 and ending at

its 3' end with the nucleotide in position 524 of SEQ ID NO 1, or a biologically active polynucleotide derivative of SEQ ID NO 2:

CTGCAGCAGGTGACGTCGTTGTTCAGCCAGGTGGGCGGCACCGGCGGCGG
CAACCCAGCCGACGAGGAAGCCGCGCAGATG
GGCCTGCTCGGCACCAGTCCGCTGTCGAACCATCCGCTGGCTGGTGGATC
AGGCCCCAGCGCGGGCGCGGGCCTGCTGCG
CGCGGAGTCGCTACCTGGCGCAGGTGGGTCGTTGACCCGCACGCCGCTGA
TGTCTCAGCTGATCGAAAAGCCGGTTGCCC
CCTCGGTGATGCCGGCGGCTGTTGCCGGATCGTCGGTGACGGGTGGCGCC
GCTCCGGTGGGTCCGGGAGCGATGGGCCAG
GGTTCGCAATCCGGCGGCTCCACCAGCCCGGGTCTGGTCGCGCCGGCACC
GCTCGCGCAGGAGCGTGAAGAAGACGACGA
GGACGACTGGGACGAAGAGGACGACTGGTGAGCTCCCGTAATGACAACA
GACTTCCCGGCCACCCGGGCCGGAAGACTTG
CCAACATTTTGGCGAGGAAGGTAAAGAGAGAAAGTAGTCCAGC

(c) a polynucleotide comprising the following nucleotide sequence of SEQ ID NO 3, starting at its 5' end with the nucleotide in position 1 of SEQ ID NO 1 and ending at its 3' end with the nucleotide in position 481 of SEQ ID NO 1, or a biologically active polynucleotide derivative of SEQ ID NO 3:

CTGCAGCAGGTGACGTCGTTGTTCAGCCAGGTGGGCGGCACCGGCGGCGG
CAACCCAGCCGACGAGGAAGCCGCGCAGATG
GGCCTGCTCGGCACCAGTCCGCTGTCGAACCATCCGCTGGCTGGTGGATC
AGGCCCCAGCGCGGGCGCGGGCCTGCTGCG
CGCGGAGTCGCTACCTGGCGCAGGTGGGTCGTTGACCCGCACGCCGCTGA
TGTCTCAGCTGATCGAAAAGCCGGTTGCCC
CCTCGGTGATGCCGGCGGCTGTTGCCGGATCGTCGGTGACGGGTGGCGCC
GCTCCGGTGGGTCCGGGAGCGATGGGCCAG
GGTTCGCAATCCGGCGGCTCCACCAGCCCGGGTCTGGTCGCGCCGGCACC
GCTCGCGCAGGAGCGTGAAGAAGACGACGA

GGACGACTGGGACGAAGAGGACGACTGGTGAGCTCCCGTAATGACAACA
GACTTCCCGGCCACCCGGGCCGGAAGACTTG

(d) a polynucleotide comprising the following nucleotide sequence of SEQ ID NO 4, starting at its 5' end with the nucleotide in position 525 of SEQ ID NO 1 and ending at its 3' end with the nucleotide in position 826 of SEQ ID NO 1 coding for the LHP polypeptide:

ATGGCAGAGATGAAGACCGATGCCGCTACCCTCGGGC
AGGAGGCAGGTAATTTCGAGCGGATCTCCGGCGACCTGAAAACCCAGATC
GACCAGGTGGAGTCGACGGCAGGTTCGTTG
CAGGGCCAGTGGCGCGGCGCGGCGGGGACGGCCGCCCAGGCCGCGGTGG
TGCGCTTCCAAGAAGCAGCCAATAAGCAGAA
GCAGGAACTCGACGAGATCTCGACGAATATTCGTCAGGCCGGCGTCCAAT
ACTCGAGGGCCGACGAGGAGCAGCAGCAGG
CGCTGTCCTCGCAAATGGGCTTCTG

(e) a polynucleotide comprising at least 12 consecutive nucleotides of a polynucleotide chosen among the group consisting of SEQ ID NO 2, SEQ ID NO 3 or SEQ ID NO 4;

(f) a polynucleotide having a sequence fully complimentary to a polynucleotide chosen among the group consisting of SEQ ID NO 2, SEQ ID NO 3 or SEQ ID NO 4;

(g) a polynucleotide hybridizing under stringent hybridization conditions with a polynucleotide chosen among the group consisting of SEQ ID NO 2, SEQ ID NO 3 or SEQ ID NO 4.

By a biologically active polynucleotide derivative of SEQ ID NO 2 or SEQ ID NO 3 according to the present invention is meant a polynucleotide comprising or alternatively consisting in a fragment of said polynucleotide which is functional as a regulatory region for expressing a recombinant polypeptide in a recombinant cell host.

More specifically, a typical biologically active polynucleotide derivative of SEQ ID NO 2 or SEQ ID NO 3 is a polynucleotide comprising at least the nucleotide region containing one transcription start site chosen among the transcription start sites respectively located at the nucleotide in position 454 of SEQ ID NO 1 and at the nucleotide in position 513 of SEQ ID NO 1.

In a particular embodiment of a biologically active derivative of SEQ ID NO 2 or SEQ ID NO 3 the ribosome binding site (shine Dalgarno sequence) which is located from the nucleotide at position 508 to the nucleotide at position 512 of SEQ ID NO 1 may be removed or absent and optionally replaced by a suitable natural or synthetic ribosome binding site, depending on the recombinant cell host in which its expression is desired.

As shown by the inventors, LHP is produced in short term culture filtrates of Mycobacterium tuberculosis, thus in the same time as ESAT-6. It is greatly expected that LHP and ESAT-6 have a synergistic action in inducing a protective immune response against a pathogenic mycobacterium, specifically mycobacteria belonging to the tuberculosis-complex. Thus, it is a preferred embodiment of the present invention to obtain a composition containing simultaneously LHP and ESAT-6, optionally in combination with other antigenic proteins from Mycobacterium tuberculosis, such as, for example, the 45/47 kDa protein or the 19 kDa, DES, ERP (28Kd) or any protein identified by biochemical or genetic means. Such a composition containing both at least LHP and ESAT-6 may be under the form of a polypeptide composition or under the form of a composition of live recombinant cell host expressing both proteins or an admixture of recombinant cell hosts each expressing one protein chosen among LHP or ESAT-6, the whole compositions being useful for immunodiagnostics or vaccine purposes.

In a specific embodiment of a recombinant vector according to the present invention, such a recombinant vector contains a regulatory polynucleotide of the invention which is placed in the suitable frame with regards to a polynucleotide containing two open reading frames encoding respectively LHP and ESAT-6. Such a plasmid may

be, for example, pIPX26 that has been deposited at the CNCM under the Accession Number I-1706 (see Figure 1). Another suitable recombinant plasmid is plasmid pPX1 that is contained in the E. coli strain that has been deposited at the CNCM on May 14, 1996, under the Accession Number I-1707 (see Figure 11).

In order to identify the relevant biologically active polynucleotide derivatives of the invention that are described hereinbefore, the one skilled in the art will refer to the Example 5 and 6 of the instant specification in order to use a recombinant vector carrying a marker gene the expression of which will be detected when placed under the control of a biologically active derivative polynucleotide of SEQ ID NO 2 or 3.

Said regulatory polynucleotides may be prepared from any of the SEQ ID NO 1, SEQ ID NO 2 or SEQ ID NO 3 by cleavage using the suitable restriction enzymes.

Said regulatory polynucleotides may also be prepared by digestion of any of SEQ ID NO 1, SEQ ID NO 2 or SEQ ID NO 3 by an exonuclease enzyme, such as for example Bal31 (Wabiko et al., 1986).

Another object of the present invention is a recombinant vector containing a polynucleotide of SEQ ID NO 2 or SEQ ID NO 3, or a biologically polynucleotide derivative thereof, and a polynucleotide coding for a polypeptide.

In a specific embodiment of the recombinant vector according to the present invention, the polynucleotide of SEQ ID NO 2 or one of its biologically active derivatives, or a biologically active derivative of SEQ ID NO 3 lacking the ribosome binding site sequence will have to be located in the suitable frame with an heterologous Shine-Dalgarno type sequence in order to allow the expression of the polypeptide encoding gene placed under its control.

The preferred expression vectors carrying the polynucleotide of SEQ ID NO 2 or SEQ ID NO 3 or one of their biologically active polynucleotide derivatives are the conventional vectors used for polypeptide expression in bacteria, such as for example plasmids of the pUC family or plasmids of the pAL family.

A specific recombinant vector according to the present invention is the plasmid pIPX30 which has been deposited at the CNCM on February 13, 1997 under the Accession Number I-1845. A map of plasmid pIPX30 is represented on Figure 7.

The polypeptide encoded by a polynucleotide contained in a recombinant vector according to the present invention may be any kind of polypeptide either of eukaryotic or prokaryotic origin.

Preferably said polynucleotide codes for an antigenic protein of a mycobacterium, and preferably a mycobacterium belonging to the Mycobacterium tuberculosis complex.

In a most preferred embodiment, the encoded antigenic polypeptide or protein is a polypeptide which undergoes post translational modifications in the mycobacterium, such as phosphorylation, glycosylation or acylation. Such preferred postranslationally modified antigenic mycobacterial polypeptides are, for example, the 19 kDa antigen from Mycobacterium tuberculosis, the expression of which is described by Herrmann et al., 1996, Harris et al., 1994 and by Garbe et al., 1993, and possibly LHP or ESAT-6.

Other antigenic mycobacterial polypeptides of interest that may be expressed under the control of a regulatory polynucleotide according to the present invention are the following: DnaK, GroEL, GroES, the 45/47 kD, ERP, and DES polypeptides from Mycobacterium tuberculosis (Bengard et al., 1994).

The present invention concerns also the polynucleotide insert of a recombinant vector as defined hereinbefore.

The invention also concerns a recombinant cell host containing a purified polynucleotide insert as defined hereinbefore or a recombinant vector according to the invention.

The recombinant cell host may be a bacteria, such as for example E. coli.

A recombinant cell host according to the present invention consists in a fast growing or a slow growing mycobacterium. Preferably, it consists in a mycobacterium belonging to the Mycobacterium tuberculosis complex, more specifically the species Mycobacterium tuberculosis itself or Mycobacterium bovis-BCG or mutants of these strains. Another embodiment of a mycobacterium recombinant cell host according to the present invention consists in Mycobacterium smegmatis.

Another object of the present invention consists in a purified polypeptide produced by a recombinant cell host according to the invention.

A method for preparing such a recombinant polypeptide comprises typically the steps of: (a) optionally preparing a recombinant vector as described above; (b) optionally introducing said recombinant vector in a suitable eukaryotic or prokaryotic cell host; (c) cultivating the recombinant cell host of step (b); (d) purifying the recombinant polypeptide produced in the culture supernatant medium or in the recombinant cell host cell lysate.

In another aspect of the present invention, polynucleotides of SEQ ID NO 2, SEQ ID NO 3 or SEQ ID NO 4 are useful as starting material in order to design new polynucleotides that hybridize specifically under stringent hybridization conditions with the polynucleotide of SEQ ID NO 1, said new polynucleotides being used as oligonucleotide primers or probes.

Consequently is also part of the present invention a polynucleotide or oligonucleotide comprising at least 12 consecutive nucleotides of a polynucleotide chosen among the group consisting of SEQ ID NO 2, SEQ ID NO 3 or SEQ ID NO 4.

By a polynucleotide or oligonucleotide hybridizing under stringent hybridization conditions according to the present invention is meant a polynucleotide that hybridizes with a polynucleotide of SEQ ID NO 2, SEQ ID NO 3 or SEQ ID NO 4 under the following hybridization conditions:

The hybridization step is realized at 65°C in the presence of 6 x SSC buffer, 5 x Denhardt's solution, 0.5% SDS and 100µg/ml of salmon sperm DNA.

The hybridization step is followed by four washing steps:

- two washings during 5 min, preferably at 65°C in a 2 x SSC and 0.1% SDS buffer;
- one washing during 30 min, preferably at 65°C in a 2 x SSC and 0.1% SDS buffer;
- one washing during 10 min, preferably at 65°C in a 0.1 x SSC and 0.1% SDS buffer.

Thus, the polynucleotides of SEQ ID NO 2, SEQ ID NO 3 and SEQ ID NO 4, or the nucleic fragments obtained from such polynucleotides may be used to select nucleotide primers notably for an amplification reaction such as the amplification reactions further described.

PCR is described in the U.S. Patent No 4,683,202. The amplified fragments may be identified by an agarose or a polyacrylamide gel electrophoresis, or by a capillary electrophoresis or alternatively by a chromatography technique (gel filtration, hydrophobic chromatography or ion exchange chromatography). The specificity of the amplification may be ensured by a molecular hybridization using as nucleic probes the polynucleotides SEQ ID NO 2, SEQ ID NO 3 and SEQ ID NO 4, fragments thereof, oligonucleotides that are complimentary to these polynucleotides or fragment thereof or their amplification products themselves.

Amplified nucleotide fragments are used as probes that are useful in hybridization reactions in order to detect the presence of one polynucleotide according to the present invention or in order to detect mutations in the SEQ ID NO 2, SEQ ID NO 3 and SEQ ID NO 4.

Are also part of the present invention the amplified nucleic fragments («amplicons») defined herein above.

These probes and amplicons may be radioactively or non-radioactively labeled, using for example enzymes or fluorescent compounds.

Such nucleic acid fragments may be used as pairs in order to amplify specific regions of SEQ ID NO 2, SEQ ID NO 3 and SEQ ID NO 4.

Preferred nucleic acid fragments that can serve as primers according to the present invention are the following:

SEQ ID NO 14: 5'-CTGCAGCAGGTGACGTCGTTG-3' (from nucleotide in position 1 to the nucleotide in position 21 of SEQ ID NO 1.

SEQ ID NO 15: 5'-CCGGGTGGCCGGGAAGTCTGTGT-3' (complimentary of the sequence from nucleotide in position 468 to the nucleotide in position 446 of SEQ ID NO 1).

SEQ ID NO 16: 5'-ACTACTTTCTCTTTCTACCTTCC-3' (complimentary of the sequence from nucleotide in position 519 to the nucleotide in position 497 of SEQ ID NO 1).

The above described primers are used in combination for performing a nucleic acid amplification of one polynucleotide according to the present invention. Suitable pairs of primers used are the following: (a) SEQ ID NO 14 and SEQ ID NO 15; (b) SEQ ID NO 14 and SEQ ID NO 16.

It is no need to say that any one of the above described primers may be also used as specific probes according to the invention.

The primers may also be used as oligonucleotide probes to specifically detect a polynucleotide according to the invention.

The primers may also be used as oligonucleotide probes to specifically detect a polynucleotide according to the invention.

Other techniques related to nucleic acid amplification may also be used and are generally preferred to the PCR technique.

The Strand Displacement Amplification (SDA) technique (Walker et al., 1992) is an isothermal amplification technique based on the ability of a restriction enzyme to cleave one of the strands at his recognition site (which is under a hemiphosphorothioate form) and on the property of a DNA polymerase to initiate the synthesis of a new strand from the 3'OH end generated by the restriction enzyme and on the property of this DNA polymerase to displace the previously synthesized strand being localized downstream. The SDA method comprises two main steps: (a) the synthesis in the presence of dCTP-alpha-S, of DNA molecules that are flanked by the restriction sites that may be cleaved by an appropriate enzyme; (b) the exponential amplification of these DNA molecules modified as such by enzyme cleavage, strand displacement and copying of the displaced strands. The steps of cleavage, strand displacement and copying of the displaced strands. The steps of cleavage, strand

displacement and copying are repeated a sufficient number of times in order to obtain an accurate sensitivity of the assay.

The SDA technique was initially realized using the restriction endonuclease HincII but is now generally practiced with an endonuclease from Bacillus stearothermophilus (BSOBI) and a fragment of a DNA polymerase which is devoid of and 5'Õ3' exonuclease activity isolated from Bacillus cladotenax (exo- Bca) [=exo-minus-Bca]. Both enzymes are able to operate at 60°C and the system is now optimized in order to allow the use of dUTP and the decontamination by UDG. When using this technique as described by Spargo et al. In 1996, the doubling time of the target DNA is of 26 seconds and the amplification rate is of 1010 after an incubation time of 15 min at 60°C.

The SDA amplification technique is more easy to perform than PCR (a single thermostated water bath device is necessary) and is faster than the other amplification methods.

Thus, another object of the present invention consists in using the nucleic acid fragments according to the invention (primers) in a method of DNA or RNA amplification according to the SDA technique. For performing of SDA, two pairs of primers are used: a pair of external primers (B1, B2) consisting in a sequence specific of the target polynucleotide of interest and a pair of internal primers (S1, S2) consisting in a fusion oligonucleotide carrying a site that is recognized by a restriction endonuclease, for example the enzyme BSOBI.

As an illustrative embodiment of the use of the primers according to the invention in a SDA amplification reaction, a sequence that is non specific for the target polynucleotide and carrying a restriction site for HincII or BSOBI is added at the 5' end of a primer specific either for SEQ ID NO 2, SEQ ID NO 3 and SEQ ID NO 4. Such an additional sequence containing a restriction site that is recognized by BSOBI is advantageously the following sequence: GCATCGAATGCATGTCTCGGGT, (SEQ ID NO:27) the nucleotides represented in bold characters corresponding to the recognition site of the enzyme BSOBI. Thus, primers useful for performing SDA amplification may be designed from any of the primers according to the invention as described above and

are part of the present invention. The operating conditions to perform SDA with such primers are described in Spargo et al., 1996.

More specifically, the following conditions are used when performing the SDA amplification reaction with the primers of the invention designed to contain a BSOBI restriction site: BSOBI/exo-Bca [=exo-minus-Bca] SDA reactions are performed in a 50 l volume with final concentrations of 9.5 mM MgCl2, 1.4 mM each dGTP, dATP, TTP, dCTP-alpha-S, 100 µg/ml acetylated bovine serum albumin, 10 ng/ml human placental DNA, 35 mM K2HPO4 pH 7.6, 0.5 µM primers S1 BSOBI and B2 BSOBI, 0.05 µM primers B1 BSOBI and B2 BSOBI, 3.2 U/µl BSOBI enzyme, 0.16 U/µl exo-Bca [=exo-minus-Bca] enzyme, 3 mM Tris-HCl, 11 mM NaCl, 0.3 mM DTT, 4 mM KCl, 4% glycerol, 0.008 mM EDTA, and varying amounts of target DNA. Prior to the addition of BSOBI and exo-Bca, incomplete reactions (35 µl) are heated at 95°C for 3 min to denature the target DNA, followed by 3 min at 60°C to anneal the primers. Following the addition of a 15 µl enzyme mix consisting of 4 µl of BSOBI (40 Units/µl), 0.36 µl exo-Bca (22 Units/µl), and 10.6 µl enzyme dilution buffer (10 mM Tris HCl, 10 mM MgCl2, 50 mM NaCl, 1 mM DTT), the reactions are incubated at 60°C for 15 min. Amplification is terminated by heating for 5 min in a boiling water bath. A no-SDA sample is created by heating a sample in a boiling water bath immediately after enzyme addition. Aerosol resistant tips from Continental Laboratory Products are used to reduce contamination of SDA reactions with previously amplified products.

The polynucleotides of SEQ ID NO 2, SEQ ID NO 3 and SEQ ID NO 4 and their above described fragments, especially the primers according to the invention, are useful as technical means for performing different target nucleic acid amplification methods such as:

- TAS (Transcription-based Amplification system), described by Kwoh et al. in 1989;
- SR (Self-sustained Sequence Replication), described by Guatelli et al. in 1990;
- NASBA (Nucleic Acid Sequence Based Amplification), described by Kievitis et al. in 1991.

- TMA (Transcription Mediated Amplification).

The polynucleotides of SEQ ID NO 2, SEQ ID NO 3 and SEQ ID NO 4 and their above described fragments, especially the primers according to the invention, are also useful as technical means for performing methods for amplification or modification of a nucleic acid used as a probe, such as:

- LCR (Ligase Chain Reaction), described by Landegren et al. in 1988 and improved by Barney et al. in 1991 who employ a thermostable ligase.

- RCR (Repair Chain Reaction), described by Segev et al. in 1992.

- CPR (Cycling Probe Reaction), described by Duck et al. in 1990.

- Q-Beta Replicase Reaction, described by Miele et al. in 1983 and improved by Chu et al. in 1986, Lizardi et al. in 1988 and by Burg et al. and Stone et al. in 1996.

When the target polynucleotide to be detected is a RNA, for example, a mRNA, a reverse transcriptase enzyme will be used before the amplification reaction in order to obtain a cDNA from the RNA contained in the biological sample. The generated cDNA is subsequently used as the nucleic acid target for the primers or the probes used in an amplification process or a detection process according to the present invention.

Thus, another object of the present invention consists in a method for detecting Mycobacterium tuberculosis in a biological sample comprising the steps of: (a) bringing into contact the nucleic acid molecules contained in the biological sample with a pair of purified polynucleotides primers derived from a polynucleotide of SEQ ID NO 2, SEQ ID NO 3 or SEQ ID NO 4; (b) amplifying said nucleic acid molecules; (c) detecting the nucleic acid fragments that have been amplified, for example, by gel electrophoresis or with a labeled polynucleotide hybridizing specifically with a polynucleotide of SEQ ID NO 2, SEQ ID NO 3 or SEQ ID NO 4.

The invention concerns also the above method, wherein before step (a), the nucleic acid molecules of the biological sample have been made available to a hybridization reaction.

The invention is also related to a kit for detecting a Mycobacterium tuberculosis bacterium in a biological sample comprising: (a) a pair of purified oligonucleotides primers according to the invention; (b) reagents necessary to perform a nucleic acid

amplification reaction; (c) optionally, a purified polynucleotide according to anyone of claims useful as a probe.

The non-labeled polynucleotides or oligonucleotides of the invention may be directly used as probes. Nevertheless, the polynucleotides or oligonucleotides are generally labeled with a radioactive element (^{32}P, ^{35}S, ^{3}H, ^{125}I) or by a non-isotopic molecule (for example, biotin, acetylaminofluorene, digoxigenin, 5-bromodesoxyuridin, fluorescein) in order to generate probes that are useful for numerous applications.

Examples of non-radioactive labeling of nucleic acid fragments are described in the French Patent No FR-7810975 or by Urdea et al. or Sanchez-Pescador et al., 1988. In the latter case, other labeling techniques may be also used such those described in the French Patent Nos. FR-2,422,956 and 2,518,755. The hybridization step may be performed in different ways (Matthews et al., 1988). The more general method consists in immobilizing the nucleic acid that has been extracted from the biological sample on a substrate (nitrocellulose, nylon, polystyrene) and then to incubate, in defined conditions, the target nucleic acid with the probe. Subsequently to the hybridization step, the excess amount of the specific probe is discarded and the hybrid molecules formed are detected by an appropriate method (radioactivity, fluorescence or enzyme activity measurement).

Advantageously, the probes according to the present invention may have structural characteristics such that they allow the signal amplification, such structural characteristics being, for example, branched DNA probes as those described by Urdea et al. in 1991 or in the European Patent No. EP-0225,807 (Chiron).

In another advantageous embodiment of the probes according to the present invention, the later may be used as «capture probes», and are for this purpose immobilized on a substrate in order to capture the target nucleic acid contained in a biological sample. The captured target nucleic acid is subsequently detected with a second probe which recognizes a sequence of the target nucleic acid which is different from the sequence recognized by the capture probe.

The oligonucleotide fragments useful as probes or primers according to the present invention may be prepared by cleavage of the polynucleotides of SEQ ID NO 2, SEQ ID NO 3 and SEQ ID NO 4 by restriction enzymes.

The experimental procedure conditions suitable for using the restriction enzymes are described in Sambrook et al. (1989).

Another appropriate preparation process of the nucleic acids of the invention containing at most 200 nucleotides (or 200 bp if these molecules are double stranded) comprises the following steps:

- synthesizing DNA using the automated method of beta-cyanethylphosphoramidite described in 1986;
- cloning the thus obtained nucleic acids in an appropriate vector;
- purifying the nucleic acid by hybridizing an appropriate probe according to the present invention.

A chemical method for producing the nucleic acids according to the invention which have a length of more than 200 nucleotides (or 200 bp if these molecules are double stranded) comprises the following steps:

- assembling the chemically synthesized oligonucleotides, having different restriction sites at each end;
- cloning the thus obtained nucleic acids in an appropriate vector;
- purifying the nucleic acid by hybridizing an appropriate probe according to the present invention.

In the case in which the above nucleic acids are used as coding sequences in order to produce a polypeptide according to the present invention, it is important to ensure that their sequences are compatible (in the appropriate reading frame) with the amino acid sequence of the polypeptide to be produced.

The oligonucleotide probes according to the present invention may also be used in a detection device comprising a matrix library of probes immobilized on a substrate, the sequence of each probe of a given length being localized in a shift of one or several bases, one from the other, each probe of the matrix library thus being complimentary of a distinct sequence of the target nucleic acid. Optionally, the substrate of the matrix may be a material able to act as an electron donor, the detection of the matrix positions in which an hybridization has occurred being subsequently determined by an electronic device. Such matrix libraries of probes and methods of specific detection of a target nucleic acid is described in the

European Patent Application No. EP-0713,016 (Affymax Technologies) and also in the U.S. Patent No. US-5,202,231 (Drmanac).

Thus, another object of the present invention consists in a method for detecting the presence of Mycobacterium tuberculosis bacteria in a biological sample comprising the steps of: (a) bringing into contact a purified polynucleotide derived from SEQ ID NO 2, SEQ ID NO 3 or SEQ ID NO 4 with a nucleic acid contained in the biological sample; (b) detecting the hybrid nucleic acid molecule formed between said purified polynucleotide and the nucleic acid molecules contained within the biological sample.

In a particular embodiment of the above method, the nucleic acid molecules of the biological sample have been made available to a hybridization reaction before performing step (a).

The invention also concerns a method for detecting a Mycobacterium tuberculosis bacterium in a biological sample comprising the steps of: (a) bringing into contact a purified polynucleotide probe according to the invention that has been immobilized onto a substrate with a biological sample; (b) bringing into contact the hybrid nucleic acid molecule formed between said purified polynucleotide and the nucleic acid contained in the biological sample with a labeled polynucleotide probe according to the invention, provided that the probe of step (a) and the probe of step (b) have non-overlapping nucleotide sequences.

The invention pertains also to the above method wherein, before step (a), the nucleic acid molecules of the biological sample have been made available to a hybridization reaction.

The invention is also directed to the above method wherein, before step (b), the nucleic acid molecules that are not hybridized with the immobilized purified polynucleotide are removed.

Another object of the present invention consists in a kit for detecting a Mycobacterium tuberculosis bacterium genus in a biological sample comprising; (a) a purified polynucleotide probe according to the invention; (b) reagents necessary to perform a nucleic acid hybridization reaction.

The invention also pertains to a kit for detecting a Mycobacterium tuberculosis bacterium in a biological sample comprising: (a) a purified polynucleotide probe

according to the invention that is immobilized onto a substrate; (b) reagents necessary to perform a nucleic acid hybridization reaction; (c) a purified polynucleotide probe according to the invention which is radioactively or non-radioactively labeled, provided that the probe of step (a) and the probe of step (b) have non-overlapping nucleotide sequences.

As already specified, the present inventors have characterized a new polypeptide, named LHP, that is encoded by the polynucleotide sequence of SEQ ID NO 1, and more precisely by the polynucleotide of sequence SEQ ID NO 4. The polynucleotide of SEQ ID NO 4 encodes the LHP polypeptide of SEQ ID NO 5 which is described hereunder.

Thus, another object of the present invention consists in a purified polypeptide, named LHP, and having the following amino acid sequence SEQ ID NO 5:

MAEMKTDAATLGQEAGNFERISGDLKTQIDQVESTAGSLQGQWRGAAGTAA
QAAVVRFQEAANKQKQELDEISTNIRQAGVQYSRADEEQQQALSSQMGF

The correspondence between the one letter-code and the three letter-codes for amino acids is found in the book of Stryer Biochemistry, Third Ed. (1988), which is incorporated here by reference for all purposes.

In both immunodiagnostics and vaccine preparation it is often possible and practical to prepare antigens from segments of a known immunogenic protein or polypeptide. Certain epitopic regions may be used to produce responses similar to those produced by the entire antigenic polypeptide. Potential antigenic or immunogenic regions may be identified by any of a number of approaches, e.g., Jameson-Wolf or Kyte-Doolittle antigenicity analysis or Hopp and Woods (1981) hydrophobicity analysis (see e.g., Jameson-Wolf, 1988; Kyte and Doolittle, 1982; U.S. Patent No. 4,554,101). Hydrophobicity analysis assigns average hydophilicity values to each amino acid residue from these values average hydrophilicities can be calculated and regions of greatest hydrophilicity determined. Using one or more of these methods,

regions of predicted antigenicity are derived from the amino acid sequence assigned to the polypeptides according to the present invention.

The present invention is also directed to portions of the polypeptide of amino acid sequence of SEQ D NO 5 that are highly immunogenic and which may thus serve as components of an immunogenic composition or a vaccine composition for the purpose of diagnosing or preventing an Mycobacterium tuberculosis infection in a patient.

In order to identify the relevant antigenic or immunogenic portions of the polypeptide of SEQ ID NO 5, one skilled in the art may bring a specific peptide derived from the polypeptide of SEQ ID NO 5 in the presence of a serum sample of a patient infected with Mycobacterium tuberculosis and then detect the complex eventually formed between the antibodies contained in the serum sample and the peptide being assayed. Such a screening assay used to define the relevant immunogenic portions of the polypeptide of SEQ ID NO 5 is advantageously a conventional ELIZA type assay, wherein, as an illustrated embodiment, radioactively or fluorescently anti-Ig antibodies are used for detecting the antigen-antibody complexes formed.

Antigenic portions of the LHP polypeptide may be obtained by enzymatic cleavage of the parent purified polypeptide, one skilled in the art being guided by the digestion map of the polypeptide of SEQ ID NO 5.
Preferred antigenic portion of the polypeptide according to the present invention are comprising the hydrophilic parts of the LHP polypeptide.

Thus, the preferred antigenic portions of a polypeptide according to the invention comprise peptides or pseudopeptides derived from the following peptides consisting in: (a) amino acid in position 1 to amino acid in position 48 of SEQ ID NO 5; (b) amino acid in position 60 to amino acid in position 100 of SEQ ID NO 5; which represent the most hydrophilic regions of the LHP polypeptide of the invention.

Specific immunogenic portions of the polypeptide of SEQ ID NO 5 characterized by the inventors are the following:

(a) SEQ ID NO 6:

NH2-MAEMKTDAATLGQEAGNFERISGDLKTQIDQVESTAGSLQGQ WRGAAGT-COOH;

(b) SEQ ID NO 7: NH2-QEAANKQKQELDEISTNIRQAGVQYSRADEEQQQ ALSSQMGF-COOH;

(c) SEQ ID NO 8: NH2-QEAGNFERISGDLKTQIDQV-COOH;

(d) SEQ ID NO 9: NH2-GDLKTQIDQVESTAGS-COOH;

(e) SEQ ID NO 10: NH2-GSLQGQWRGAAGTAAA-COOH;

(f) SEQ ID NO 11: NH2-QEAANKQKQELDEIST-COOH;

(g) SEQ ID NO 12: NH2-STNIRQAGVQYSRADEEQQQALSSQMGF-COOH;

(h) SEQ ID NO 13: NH2-RADEEQQQALSSQMGF-COOH.

In a preferred embodiment of the immunogenic polypeptide according to the present invention, the epitope unit of said polypeptide have from 6 to 50 amino acids in length, preferably from 6 to 20 amino acids in length and most preferably from 6 to 15 amino acids in length, and is capable to induce in vivo a protective immune response against the LHP antigen which is expressed by Mycobacterium tuberculosis. An immunogenic polypeptide having a long amino acid chain (from 25 to 50 amino acids in length) is preferably used in case of conformational epitope units. Furthermore, a large epitope unit is expected to carry both a B-epitope and a T-epitope.

By an epitope or an epitope unit according to the present invention is meant a portion of the LHP polypeptide which is delinated by the area of interaction with antibodies that are specific to LHP, in particular monoclonal antibodies directed against LHP. The above disclosed immunogenic portions of the LHP polypeptide of SEQ ID NO 5 are all bearing at least one epitope unit.

Are also part of the immunogenic polypeptides of the present invention those polypeptides which comprise, but are not limited to, at least one epitope unit

recognized by a monoclonal antibody directed against the LHP polypeptide or a peptide fragment thereof.

Specifically, the monoclonal or polyclonal antibody according to the invention recognizes the LHP polypeptide of SEQ ID NO 5 or one peptide fragment thereof.

The antibodies may be prepared from hybridomas according to the technique described by Phalipon et al. in 1995 or also by Kohler and Milstein in 1975. The polyclonal antibodies may be prepared by immunization of a mammal, especially a mouse or a rabbit, with a polypeptide according to the invention that is combined with an adjuvant of immunity, and then by purifying of the specific antibodies contained in the serum of the immunized animal on an affinity chromatography column on which has previously been immobilized the polypeptide that has been used as the antigen.

The present invention is also directed to a diagnostic method for detecting the presence of a Mycobacterium tuberculosis is a biological sample, said diagnostic method comprising the steps of: (a) bringing into contact the biological sample expected to contain a Mycobacterium tuberculosis bacterium with a purified monoclonal or polyclonal antibody according to the invention; (b) detecting the antigen-antibody complexes formed.

In a specific embodiment of the above diagnostic method, step (a) is preceded by preparing a purified preparation of the said anti-immunogenic polypeptide monoclonal or polyclonal antibody.

In a preferred embodiment of the above diagnostic method, said method consists in an immunoassay including enzyme linked immunoassay (ELIZA), immunoblot techniques, as well as radio-immunoassays (RIA) which preceding techniques are all available from the prior art.

A typical preferred immunoassay according to the invention comprises the following steps: (a) incubating microtitration plate wells with increasing dilutions of the

biological sample to be assayed; (b) introducing in said microtitration plate wells with a given concentration of a monoclonal or polyclonal antibody according to the invention; (c) adding a labeled antibody directed against human or animal immunoglobulins, the labeling of said antibodies being, for example, an enzyme that is able to hydrolyze a substrate molecule, the substrate molecule hydrolysis inducing a change in the light absorption properties of said substrate molecule at a given wavelength, for example at 550 nm.

The present invention also concerns a diagnostic kit for the in vitro diagnosis of an infection by Mycobacterium tuberculosis, comprising the following elements: (a) a purified preparation of a monoclonal or a polyclonal antibody according to the invention; (b) suitable reagents allowing the detection of the antigen/antibody complexes formed, these reagents preferably carrying a label compound (a marker), or being recognized themselves by a labeled reagent; (c) optionally, a reference biological sample containing the pathogenic microorganism antigen recognized by the purified monoclonal or polyclonal antibody (positive control); (d) optionally, a reference biological sample that does not contain the pathogenic microorganism antigen recognized by the purified monoclonal or polyclonal antibody (negative control).

The present invention is also directed to a polyclonal or a monoclonal antibody directed against an immunogenic peptide according to the invention.

Are also part of the present invention polypeptides that are homologous to the initially selected polypeptide bearing at least an epitope unit. By homologous peptide according to the present invention is meant a polypeptide containing one or several amino acid substitutions in the amino acid sequence of the initially selected polypeptide carrying an epitope unit. In the case of an amino acid substitution, one or several - consecutive or non-consecutive- amino acids are replaced by «equivalent» amino acids. The expression «equivalent» amino acid is used herein to name any amino acid that may substituted for one of the amino acids belonging to the initial polypeptide structure without decreasing the binding properties of the corresponding

peptides to the monoclonal antibody that has been used to select the parent peptide and without decreasing the immunogenic properties against the specified pathogenic microorganism. Thus, an homologous polypeptide according to the present invention has the same immunological characteristics as the parent polypeptide (for example as the polypeptide of SEQ ID NO 5) with respect to the ability to confer increases resistance to infection with bacteria belonging to the tuberculosis complex.

These equivalent aminoacyles may be determined either by their structural homology with the initial aminoacyles to be replaced, by the similarity of their net charge, and by the results of the cross-immunogenicity between the parent peptides and their modified counterparts.

The peptides containing one or several «equivalent» amino acids must retain their specificity and affinity properties to the biological targets of the parent protein, as it can be assessed by a ligand binding assay or an ELIZA assay.

For example, amino acids may be placed in the following classes: non-polar, uncharged polar, basic, and acidic. Conservative substitutions whereby an amino acid of one class is replaced with another amino acid of the same type fall within the scope of the subject invention so long as the substitution does not materially alter the biological activity of the compound. Table 1 provides a listing of examples of amino acids belonging to each class.

Table 1: The Different Classes of Amino Acids

Class of Amino Acid	Examples of Amino Acids
Non Polar	A, V, L, I, P, G, F, W
Uncharged Polar	M, S, T, C, Y, N, Q
Acidic	D, E

Basic K, R, H

By modified amino acid according to the present invention is also meant the replacement of a residue in the L-form by a residue in the D-form or the replacement of a Glutamic acid (E) residue by a Pyro-glutamic acid compound. The synthesis of peptides containing at least one residue in the D-form is, for example, described by Koch et al. in 1977.

As an illustrative example, it should be mentioned the possibility to realize substitutions without a deep change in the immunogenic polypeptide binding properties of the correspondent modified peptides by replacing, for example, leucine by valine, it being understood that the reverse substitutions are permitted in the same conditions.

In order to design peptides homologous to the immunogenic polypeptides according to the present invention, one skilled in the art can also refer to the teachings of Bowie et al. (1990).

A specific, but not limitative, embodiment of a modified peptide molecule of interest according to the present invention, which consists in a peptide molecule, named herein also «pseudopeptide», which is resistant to proteolysis, is a peptide in which the -CONH- peptide bound is modified and replaced by a (CH2NH) reduced bound, a (NHCO) retro inverso bound, a (CH2-O) methylene-oxy bound, a (CH2-S) thiomethylene bound, a (CH2CH2) carba bound, a (CO-CH2) cetomethylene bound, a (CHOH-CH2) hydroxyethylene bound), a (N-N) bound, a E-alcene bound or also a -CH=CH- bound.

The immunogenic polypeptides according to the present invention may be prepared in a conventional manner by peptide synthesis in liquid or solid phase by successive coupling of the different amino acid residues to be incorporated (from the N-terminal end to the C-terminal end in liquid phase, or from the C-terminal end to the N-

terminal end in solid phase) wherein the N-terminal ends and the reactive side chains are previously blocked by conventional groups.

For solid phase synthesis the technique described by Merrifield may be used in particular. Alternatively, the technique described by Houbenweyl in 1974 may also be used or generally any chemical synthesis method well known by one skilled in the art, such as for example a chemical synthesis method performed with a device apparatus commercialized by the Applied Biosystems firm.

In order to produce a peptide chain using the Merrifield process, a highly porous resin polymer is used, on which the first C-terminal amino acid of the chain is fixed. This amino acid is fixed to the resin by means of its carboxyl groups and its amine function is protected, for example, by the t-butyloxycarbonyl group.

A peptide or pseudopeptide according to the present invention is advantageously combined with or contained in an heterologous structure, or polymerized in such a manner as to enhance their ability to induce a protective immune response against the pathogenic microorganism.

As a particular embodiment of the immunogenic polypeptide according to the present invention, said immunogenic polypeptide comprise more than one epitope unit, preferably from 2 to 20 epitope units, more preferably from 2 to 15 epitope units and most preferably 3 to 8 epitope units per polypeptide molecule, usable as an active principle of a vaccine composition.

The immunogenic polypeptides of the invention that comprise more than one epitope unit are herein termed «oligomeric polypeptides». The said polymers may be obtained by the technique of Merrifield or any other conventional peptide polymer synthesis method well known by one skilled in the art.

The peptides thus obtained may be purified, for example by high performance liquid chromatography, such as reverse phase and/or cationic exchange HPLC, as described by Rougeot et al. in 1994.

As another particular embodiment of the oligomeric immunogenic polypeptides according to the present invention, the peptides or pseudopeptides are embedded within a peptidic synthetic matrix in order to form a MAP (Multi-branched Associated Peptide) type structure. Such MAP structures as well as their method of preparation are described by Tam in 1988 or in the PCT Patent Application No. WO94/28915 (Hovanessian et al.). The embedding of the peptides or pseudopeptides of therapeutic value according to the present invention within MAP type structures are expected to cause an increase in the immunogenic and/or protective properties of the initial molecules as regards to the pathogenic microorganism infection.

In order to improve the antigenic presentation of the immunogenic polypeptides according to the present invention to the immune system, said immunogenic polypeptides are presented via a MAP (Multiple Antigen Peptide) construct. This kind of presentation system is able to present more than one copy of a selected epitope unit per molecule (4 to 8 immunogenic polypeptide mimic per MAP construct molecule) embedded in a non immunogenic «carrier» molecule.

Thus, another object of the present invention consists in peptide constructs that are able to ensure an optimal presentation of the LHP immunogenic portions of the invention to the immune system.

In a specific embodiment of the peptide constructs according to the invention, the immunogenic polypeptides (the epitope units) are part of a MAP construct as defined above, such MAP construct comprising from four to eight epitope units per molecule, for example grafted on a lysine core.

Generally, an immunogenic polypeptide according to the present invention will comprise an additional T-epitope that is covalently or non-covalently combined with said polypeptide of the invention. In a preferred embodiment, the additional T-epitope is covalently linked to the immunogenic polypeptide.

Illustrative embodiments of a suitable T-cell epitope to be combined with an immunogenic peptide mimic according to the invention are, for example, the following:

- hepatitis delta T-cell epitopes (Nisini et al., 1997);

- a T-cell epitope from the Influenza virus (Fitzmaurice et al., 1996);
- a T-cell epitope of woodchuck hepatitis virus (Menne et al., 1997);
- a T-cell epitope from the rotavirus VP6 protein (Banos et al., 1997)
- a T-cell epitope from the structural proteins of entroviruses, specifically from the VP2, VP3 and VP1 capsid proteins (Cello et al., 1996);
- a T-cell epitope from Streptococcus mutans (Senpuku et al., 1996); or also
- a T-cell epitope from the VP1 capsid protein of the foot and mouth disease virus (Zamorano et al., 1995);

Preferred additional T-epitopes used according to the present invention are for example universal T-epitopes, such as tetanus toxoid or also the VP1 poliovirus capsid protein (Graham et al., 1993).
In a most preferred embodiment, the T-cell epitope used consists in a peptide comprised between amino acid in position 103 and amino acid in position 115 of the VP1 poliovirus capsid protein.

Thus, the MAP construct may comprise an additional T-epitope which is covalently linked to the immunogenic polypeptide of the MAP, the orientation being chosen depending on the immunogenic polypeptide to be used to prepare the MAP construct. Accordingly, the additional T-epitope may be located at the external end (opposite to the core) of the MAP or conversely, the additional T-epitope may be directly linked to the core of the MAP construct, thus preceding the immunogenic polypeptide which is then external to the MAP construct.

In another embodiment of the peptide constructs according to the present invention, the immunogenic polypeptide is directly coupled with a carrier molecule such as KLH (Keyhole Limpet Hemocyanin) or preferably with tetanus toxoid.
The immunogenic polypeptide according to the invention may be presented in different additional ways to the immune system.

In one specific embodiment the immunogenic polypeptide of the invention may be presented under the form of ISCOMs (Immunostimulating complexes) that are composed of Quil A (a saponin extract from Quilaja saponaria olina bark), cholesterol and phospholpids associated with the immunogenic polypeptide (Mowat et al., 1991; Morein, 1990; Kersten et al., 1995).

The immunogenic polypeptides of the invention may also be presented under the form of biodegradable microparticles (microcapsules or microspheres) such as for example lactic and glutamic acid polymers as described by Aguado et al. in 1992, also termed Poly(lactide-co-glycolide) microcapsules or microspheres.

Other microparticles used to present the LHP-derived polypeptide antigens of the invention are synthetic polymer microparticles carrying on their surface one or more immunogenic polypeptides covalently bonded to the material of the microparticles, said immunogenic polypeptide(s) each carrying one or more epitope units and being present at a density of between 104 and 5.105 molecules / m2. These microparticles have an average diameter of between about 0.25 m and 1.5 m, and preferentially of about 1 m so as to be able to be presented to CD4+ T lymphocytes by phagocytic cells. Said microparticles are more particularly characterized in that the covalent bond is formed by reaction between the NH2 and/or CO groups of the immunogenic peptide mimic and the material making up the microparticle. Advantageously, such bond is created by bridging reagent as intermediate, such as for example glutaraldehyde or carbodiimide. The material of the microparticle can advantageously be a biocompatible polymer, such as acrylic polymer, for example polyacrolein or polystyrene or the poly)alpha-hydroxy acids), copolymers of lactic and glycolide acids or lactic acid polymers, said polymers being a homopolymer or hetero- or co-polymer. The above described microparticles characteristics are found in the French Patent Application No. FR 92-10,879 filed on September 11, 1992 (Leclerc et al).

The immunogenic polypeptide of the invention may also be included within or absorbed onto liposomes particles, such as those described in the PCT Patent Application No. PCT/FR 95/00215 published on August 31, 1995 (Riveau et al.).

The present invention is also directed to an immunogenic composition comprising an immunogenic polypeptide according to the invention, notably under the form of a MAP construct or a peptide construct as defined above, and including the oligomeric immunogenic polypeptides described hereinbefore, or also under a microparticle preparation.

The invention also pertains to a vaccine composition for immunizing human and mammal animals against a Mycobacterium tuberculosis infection, comprising an immunogenic composition as described above in combination with a pharmaceutically compatible excipient (such as saline buffer), optionally in combination with at least one adjuvant of the immunity such as aluminium hydroxide or a compound belonging to the muramyl peptide family.

A vaccine according to the present invention is preferably one which is capable of inducing a substantial and specific acquired immune resistance in a mouse or guinea pig against tuberculosis caused by mycobacteria belonging to the tuberculosis-complex, which acquired immune resistance corresponds to at least 20% of the protective immune resistance elicited by Mycobacterium bovis-BCG, as assessed by the observed reduction in mycobacterial counts from spleen, lung or other organ homogenates isolated from the mouse or guinea pig receiving a challenge infection with a virulent strain of M. tuberculosis.

The preferred acquired immune resistance corresponds to at least 50% of the protective immune response elicited by M. bovis-BCG, such as at least 60%, or even more preferred to at least 80% of the protective immune resistance elicited by M. bovis-BCG, such as at least 90% and advantageously 100%.

Various methods of achieving adjuvant effect for the vaccine include the use of agents such as aluminium hydroxide or phosphate (alun), commonly used as 0.05 to

0.1 percent solution in phosphate buffered saline, admixture with synthetic polymers of sugars (Carbopol) used as 0.25% solution. Another suitable adjuvant compounds consist in DDA (dimethyldioctadecylammonium bromide), as well as immune modulating substances, such as lymphokines (e.g. gamma-IFN, IL-1, IL-2 and IL-12) or also gamma-IFN inducers compounds, such as poly I:C.

Preparation of vaccines which contain polypeptides as active ingredients is generally well understood in the art, as exemplified by U.S. Patents 4,608,251; 4,601,903; 4,599,231; 4,599,230; 4,596,792 and 4,578,770, all incorporated herein by reference.

The vaccine composition according to the present invention is advantageously prepared as injectable either as liquid solution or suspension; solid forms suitable for solution in or suspension in, liquid prior injection may also be prepared.

The active immunogenic polypeptide contained in the vaccinal composition is generally mixed with excipients which are pharmaceutically acceptable and compatible, such as for example, water saline, dextrose, glycerol, ethanol, or a combination of more than one of the above excipients.

In addition, if desired, the vaccine composition may contain minor amounts of auxiliary substances such as wetting or emulsifying agents, pH buffering agents, or adjuvants which enhance the effectiveness of the vaccines.

The vaccines are conventionally administered parentally, by injection, for example, either subcutaneously or intramuscularly. Additional formulations are suitable for other modes of administration include suppositories and, in some cases, oral formulations, which may be preferred embodiments for the development of a desired mucosal immunity.

The immunogenic polypeptide of the invention may be formulated into the vaccine as neutral or salt forms. Pharmaceutically acceptable salts include acid addition salts (formed with free amino groups of the peptide) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic oxalic, tartaric or mandelic acid. Salts formed with the free carboxyl

groups may also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, histidine or procaine.

The vaccine compositions of the invention are administered in a manner compatible with the dosage formulation, and in such amounts as will be therapeutically effective and immunogenic. The quantity to be administered depends on the subject to be treated, including, e.g., the capacity of the individual's immune system to mount an immune response.

Suitable dosage ranges are of the order of several hundred micrograms active immunogenic polypeptide with a preferred range from about 0.1 μg to 1000 μg, such as in the range from about 1 μg to 300 μg, and especially in the range from about 10 μg to 50 μg.

The dosage of the vaccine will depend on the route of administration and will vary according to the age of the patient to be vaccinated and, to a lesser degree, the size of the person to be vaccinated.

Preferably, both in the case of an immunogenic polypeptide carrying a single epitope unit and in the case of an immunogenic polypeptide carrying several epitope units, the vaccine composition is administered to humans in the range from 0.1 to 1 μg immunogenic polypeptide per kilogram patient's body weight, preferably in the range from 0.5 μg/kg of body weight, this representing a single vaccinal dose for a given administration.

In the case of patients affected with immunological disorders, such as, for example, immunodepressed patients, each injected dose preferably contains half the weight quantity of the immunogenic polypeptide contained in a dose for a healthy patient.

In many instances, it will be necessary to proceed with multiple administrations of the vaccine composition according to the present invention, usually not exceeding six administrations, more usually not exceeding four vaccinations, and preferably one or more, usually at least about three administrations. The administrations will normally be at from two to twelve week intervals, more usually from three to five week

intervals. Periodic boosters at intervals of 1-5 years, usually three years, will be desirable to maintain the desired levels of protective immunity.

Preferably, the vaccine composition is administered several times. As an illustrative example, three vaccinal doses as defined herein above are respectively administered to the patient at time t0, at time t0 + 1 month and at time t0 + 12 months.

Alternatively, three vaccinal doses are respectively administered at time t0, at time t0 + 1 month and at time t0 + 6 months.

The course of the immunization may be followed by in vitro proliferation assays of PBL (peripheral blood lymphocytes) co-cultured with the immunogenic polypeptide of the invention, and especially by measuring the levels of gamma-IFN released from the primed lymphocytes. The assays may be performed using conventional labels, such as radionuclides, enzymes or fluorescent compounds. These techniques are well known from one skilled in the art and found notably in U.S. Patent No. 3,731,932; 4,174,384 and 3,949,064, which are herein incorporated by reference.

As described above, a measurement of the effect of the polypeptides in the vaccine compositions according to the present invention may be to assess the gamma-IFN released from memory T-lymphocytes. The stronger immune response the more gamma-IFN will be released, accordingly, a vaccine composition according to the invention comprises a polypeptide capable of releasing from the memory T-lymphocytes at least 15000 pg/ml, such as 2000 pg/ml, preferably 3000 pg/ml gamma-IFN, in the above described in vitro assays.

In mice, that are administered with a dose comparable to the dose used in humans, the antibody production is assayed after recovering the immune serum and revealing the immune complex formed between the antibodies present in the serum samples and the immunogenic polypeptide contained in the vaccine composition, using the usual methods well known from one skilled in the art.

The immunogenic polypeptides used in the vaccinal strategy according to the present invention may also be obtained using genetic engineering methods. One skilled in the art will refer to the known sequence of DNA insert that expresses a specific antigenic portion (epitope unit) of an immunogenic polypeptide of the invention and

also to the general literature to determine which appropriate codons may be used to synthesize the desired peptide.

There is no need to say that the expression of the polynucleotide that encodes the immunogenic polypeptide of interest may be optimized, according to the organism in which the sequence has to be expressed and the specific codon usage of this organism (mammal, plant, bacteria, etc.). For bacteria and plant, respectively, the general codon usages may be found in the European Patent Application No. EP-0359472 (Mycogen).

As an alternative embodiment, the epitope unit of the immunogenic polypeptide contained in a vaccine composition according to the present invention is recombinantly expressed as a part of longer polypeptide that serves as a carrier molecule.

Specifically, the polynucleotide coding for the immunogenic polypeptide of the invention, for example a polypeptide having an amino acid length between 100 and 200 amino acid residues, is inserted at least one permissive site of the polynucleotide coding for the Bordetella cyaA adenylate cyclase, for example at a nucleotide position located between amino acids 235 and 236 of the Bordetella adenylate cyclase. Such a technique is fully described in United States Patent No., 5,503,829 granted on April 2, 1996 (Leclerc et al.).

In another embodiment of the vaccine composition according to the present invention, the nucleotide sequence coding for the desired immunogenic polypeptide carrying one or more epitope units is inserted in the nucleic sequence coding for a surface protein of Haemophilus influenza, such as described in the PCT Application No. PCT/US 96/17698 (the Research Foundation of State University of New York), which is herein incorporated by reference.

In a further embodiment of the vaccine composition according to the present invention, the latter is based upon a live recombinant cell host expressing the entire LHP polypeptide of sequence SEQ ID NO 5 or alternatively a polypeptide containing an immunogenic portion of LHP according to the invention or also an oligomeric immunogenic LHP-derived polypeptide such as those described hereinbefore.

The microorganism in the vaccine may be a bacterium such as bacteria selected from the group consisting of the genera Mycobacterium, Salmonella, Pseudomonas or E. coli.

A preferred embodiment of a vaccine composition containing a live recombinant cell host according to the invention consists in a Mycobacterium bovis-BCG strain which has been transformed with a polynucleotide encoding the entire LHP polypeptide or alternatively a polypeptide containing an immunogenic portion of LHP or also an oligomeric immunogenic LHP-derived polypeptide.

An advantageous method used to transform a Mycobacterium bovis-BCG strain with a polynucleotide coding for an immunogenic polypeptide according to the present invention consists in introducing the polynucleotide of interest via an allelic exchange event (homologous recombination involving a double cross-over) or via an homologous recombination involving a single cross-over, using a recombinant vector.

Such a recombinant vector carries the gene encoding the immunogenic polypeptide or interest which has been introduced in a polynucleotide counterpart of a gene non essential for the growth of Mycobacterium bovis-BCG on the vector, such as for example the urease gene. Said vector carries advantageously also a conditional lethal selection marker such as SacB gene. The relevant transformation methods and vectors are fully described by Reyrat et al. (1995) or Pelicic et al. (1996).

Another embodiment of the live vaccine compositions according to the present invention consists in compositions containing live mycobacteria, and preferably live Mycobacterium bovis-BCG or mutant derived from Mycobacterium tuberculosis or Mycobacterium bovis-BCG transformed with a recombinant vector containing an antigenic protein placed under the control of a regulatory polynucleotide according to the present invention.

The live vaccine compositions of the invention are administered in a manner compatible with the dosage formulation, and in such amount as will be therapeutically effective and immunogenic. The quantity to be administered depends on the subject to be treated, including, e.g., the capacity of the individual's immune system to induce an immune response.

Suitable dosage ranges are of the order of 104 to 106 cfu (colony forming units) at an attenuated recombinant mycobacteria concentration of about 106 cfu/mg. Most preferably, the effective dose is about 105 cfu.

The dosage of the vaccine will depend on the route of administration and will vary according to the age of the patient to be vaccinated and, to a lesser degree, the size of the person to be vaccinated. Most preferably, the vaccine composition according to the present invention is administered via an intradermal route and in a single boost.

In the case of patients affected with immunological disorders, such as for example immunodepressed patients, each injected dose preferably contains half the weight quantity of the attenuated mycobacteria contained in a dose for a healthy patient.

In the case of neonates, the dose will be four times less than for an adult, and in case of young children (4-6 years old), the dose will be half the dose used for an adult healthy patient.

In some instances, it will be necessary to proceed with multiple administrations of the vaccine composition according to the present invention, usually not exceeding six administrations, more usually not exceeding four vaccinations, and preferably one or more, usually at least about three administrations. The administrations will normally be at from two to twelve week intervals, more usually from three to five week intervals. Periodic boosters at intervals of 1-5 years, usually three years, will be desirable to maintain the desired levels of protective immunity.

Immunization by DNA-based vaccines has been the object of several studies since the beginning of the 1990s. A DNA-based vaccine involves the transfer of a gene or at least a portion of a gene, by direct or indirect means, such that the protein subsequently produced acts as an antigen and induces a humoral-and/or cellular mediated immunological response.

Ulmer et al. - Science, 259: 1745-1749 [1993] obtained protection against the influenza virus by induction of the cytotoxic T-lymphocytes through injection of a plasmid coding for an influenza A nucleoprotein into the quadriceps of mice. The plasmid used carries either the Rous sarcoma virus promoter or the cytomegalo virus promoter.

Raz et al. -Proc. Natl. Acad. Sci. USA 90: 4523-4527 [1993] injected vectors comprising the Rous sarcoma virus promoter and a gene coding for interleukin-2, interleukin-4, or the β1-type transforming growth factor (TFG-β1). The humoral and cell-mediated immune response of the mice to which these plasmids have been intramuscularly administered are improved.

Wang et al. -Proc. Natl. Acad. Sci. USA 90: 4156-5160 [1993] injected a plasmid carrying a gene coding for the envelope protein of the HIV-1 virus into mice muscles. The plasmid injection was preceded by treatment with bupivacaine in the same area of the muscle. The authors demonstrated the presence of antibodies capable of neutralizing the HIV-1 virus infection. However, the DNA was injected twice a week for a total of four injections.

Davis et al. (Compte-Rendu du 28eme Congres Europeen sur le muscle, Bielefeld, Germany, 21-25 September 1992) injected plasmids carrying a luciferase or -galactosidase gene by pre-treating the muscle with sucrose or a cardiotoxin. The authors observed the expression of luciferase or β-galactosidase.

More recently, an article published in Science et Avenir (September 1993: 22-25) indicates that Whalen and Davis succeeded in immunizing mice against the hepatitis B virus by injecting pure DNA from the virus into their muscles. An initial injection of snake venom toxin, followed 5 to 10 days later by a DNA injection, is generally described. However, the authors specify that this method is not practical.

These studies were preceded by other experiments in which various DNAs were injected, in particular into muscle tissues. For example, U.S. Patent No. 5,589,466 and 5,580.859 (VICAL INC) and the International Application PCT/US90/01515 (published under No. WO/90/11092) disclose various plasmid constructions which can be injected in particular into muscle tissues for the treatment of muscular dystrophy. However, this later document specifies that DNA is preferentially injected in liposomes.

Additionally, Canadian Patent CA 362 96630 (published under No. 1,169,793) discloses the intramuscular injection of liposomes containing DNA coding, in particular, for HBs or HBc antigens. The results described in this patent mention the HBs antigen expression. The presence of anti-HBs antibodies was not investigated.

International Application PCT/FR92/00898 (published under No. WO93/06223) discloses viral vectors which can be conveyed to target cells by blood. These vectors are recognized by the cell receptors, such as the muscle cells, and can be used in the treatment of muscular dystrophy or thrombosis.

The present invention relates to a composition capable of inducing an immune response, and more particularly, an humoral or/and a cytotoxic response comprising a nucleotide sequence expressed in muscle cells. The nucleotide sequence comprises a gene or complimentary DNA coding for at least a portion of nucleotidic sequence comprised in the pIPX61 insert preferably the lhp polynucleotide coding region and a promoter and/or regulatory region allowing for the expression of the gene or complimentary DNA in the muscle cells.

The invention further relates to the vector, which serves as a vehicle for the gene or complimentary DNA coding for at least lhp polynucleotide coding region and a promoter allowing for the expression of the gene or cDNA which is administered to an individual to be immunized.

The present invention will be fully illustrated by the examples described below, although the scope of the invention cannot in any way be limited to these embodiments.

Examples

Example 1: Genetic organization upstream from the M. tuberculosis orflC gene.

To isolate potential promoter region, the inventors have cloned the 1.1 kb DNA sequence upstream from the M. tuberculosis orflC gene. A 150 bp DNA fragment covering the first half of the orflC gene was obtained by digestion of the plasmid pAA249 with EcoRI/PstI, radioactively labeled and used to probe a cosmid library of Mycobacterium tuberculosis strain H37Rv by colony hybridization (Sambrook et al., 1989). A 1.1 kb PstI restriction fragment shared by three strongly hybridizing cosmids, was transferred to pBluescript II KS + to give pIPX 61 (Figure 2). Double-stranded DNA sequencing revealed perfect nucleotide identity between the

1069 bp insert of pIPX61 and its counterpart in M. bovis RD1 (Maheiras et al., 1996). It included a 285 bp open reading frame preceded by a potential ribosomal binding site (AGAGA) in the same transcriptional orientation as orflC (Figure 4A). This ORF was designated lhp (L45 homologous protein) since its deduced product shared 40% peptide identity with the M. leprae L45 seroreactive antigen (Figure 4B). L45 seroreactive antigen is strongly recognized by sera from lepromatous leprosy patients (Sathish et al., 1990) but its function is currently unknown. lhp was not annotated in the M. bovis RD1 sequence published by Maheriras et al. (Maheiras et al., 1996) and overlaps with the 3' end of the predicted but uncharacterized orflB gene.

Example 2: Analysis of lhp- and orflC-lacZ gene fusions.

To investigate promoter activity, the inventors have constructed translational fusions between orflC, lhp and the lacZ reporter gene, orflC and lhp were inserted into promoter probe vectors of the pJEM series (Timm et al., 1994), out- or in frame with regard to lacZ. The resulting plasmids were named pIPX15, pIPX16 and pIPX46, pIPX47 respectively (Figure 2). These constructs were introduced by electroporation in M. smegmatis mc2 155 and β-galactosidase activity was assayed in bacterial cell extracts. Strong β-galactosidase activity was detected in extracts of cells carrying in frame fusions (pIPX16 and pIPX47) but not in extracts of cells carrying out-of-frame fusions (pIPX15 and pIPX46). Thus (i) lhp is expressed and translated (ii) there is a mycobacterial promoter activity somewhere in the 900 bp upstream from the orflC start codon. The differences in levels of β-galactosidase activity produced from pIPX16 and pIPX47 may reflect differences in the stability of lacZ fusion proteins. Alternatively, this may be attributable to differences in the efficiency of lhp and orflC translation signals. In that respect, a long (A+G)-rich stretch upstream from the orflC ATG and overlapping the predicted ribosomal binding site may potentially alter the translation of orflC-lacZ.

Example 3: Mapping of the lhp/orflC promoter activity.

The inventors have performed primer extension experiments to map the site(s) of lhp/orflC transcription initiation. Total RNA was extracted (Bashyam et al., 1994) from M. tuberculosis and M. smegmatis mc2155 transformed with pIPX16. By walking upstream from the start codon, we identified one major and two secondary transcriptional start sites in M. tuberculosis (Figure 5A). The sites are close together within a region of 30 bp about 430 nucleotides upstream from the orflC ATG start codon. The (A+T)-rich [TAATGA] region may correspond to the -10 hexcamer motif identified in promoters of other bacterial genera. The corresponding -35 region contains two tandem repeats of a 18 nucleotide motif extending from positions -25 to -60 (Figure 5B). The significance of this organization is unknown but may serve a regulatory function (Collado-Vides et al., 1991). The positions of the transcriptional start sites detected in M. tuberculosis are consistent with lhp and orflC being cotranscribed.

Surprisingly, the transcription start sites detected with RNA extracted from M. smegmatis mc2155 [pIPX16] differed from those in M. tuberculosis. The major M. smegmatis start site was immediately downstream from the predicted lhp ribosome binding site (Figure 4A). This suggests that the genuine lhp-orflC promoter may not be recognized and that alternative transcription signals are used in M. smegmatis. To test this, a 480 bp DNA fragment encompassing the transcription start sites identified in M. tuberculosis was inserted into the vector pJEM15, creating a transcriptional fusion with lacZ. The resulting plasmid (pIPX18) was introduced into M. bovis BCG and lacZ. The resulting plasmid (pIPX18) was introduced into M. bovis BCG and lacZ activity was detected by the appearance of blue colonies on 7H10 X-Gal indicator plates. In contrast, no activity was observed with colonies of M. smegmatis mc2155 [pIPX18] grown on the same plates. The absence of significant β-galactosidase activity in mc2155 [pIPX18] was confirmed by standard β-galactosidase assays (Figure 2). This data suggests that correct expression of lhp-orflC required facto(s) absent from or not functional in M. smegmatis.

Example 4 : lhp/orflC are organized as an operon.

To confirm that lhp and orflC belong to the same transcriptional unit, the inventors used the ESA-A probe (see Figure 2) for Northern blotting hybridization with total M. tuberculosis RNA. A strong hybridization signal migrating at about 800 bp (estimated using RNA molecular weight standards) was detected (Figure 6, lanes 1 and 2). Moreover, comparable amounts of this transcript were detected in early (day 5) and late (day 16) cultures (Figure 6, lanes 3 to 6). Consequently the lhp-orflC transcript appears to be produced in M. tuberculosis from a constitutive high level promoter, and/or alternatively, is very stable. In view of the position of the +1 sites, a transcript of 800 bp covers both lhp and orflC. Furthermore, there is a structure similar to a Rho-independent transcription terminator, 790 bp downstream from the lhp/orflC major transcriptional start site.

Example 5: Construction of the pIPX30 expression/tagging vector.

The design of the pIPX30 was carried out in two steps. First, we constructed a small sized E. coli-mycobacteria shuttle plasmid harboring a convenient multiple cloning site and a selectable marker gene conferring resistance to kanamycin. The PstI fragment from pUC4K together with the NdeI/BsaI-digested pUC18 were blunted with phage T4 polymerase (Amersham) and ligated to each other. The resulting plasmid designated pPV8 was digested by StuI and ligated to the EcorRV/HpaI fragment from pAL5000 to give pPV24. The pPV24 plasmid is a multipurpose shuttle cloning vector harboring four unique restriction sites (KpnI, BamHI, XbaI, PstI). It allows alpha complementation and blue/white selection of recombinants in E. coli.

In a second time, a fragment containing the ESAT-6 promoter region, was amplified by PCR under standard condition using XP1 () and XP2 () oligonucleotides. A linker fragment was generated by PCR-driven in vitro extension and amplification of the two overlapping oligonucleotides XL1 () and XL2 (). Since the promoter and linker fragments carried a 12 bp overlapping region, they were recombined in vitro

by PCR amplification using the XP1 and XL2 oligonucleotides. This PCR-tailored fragment was inserted in pPV24 previously digested with KpnI/PstI and blunted with T4 polymerase, resulting in the plasmid pIPX30. As presented Figure 1, the pIPX30 expression cassette is composed of (i) a Shine Dalgarno motif functional in mycobacteria (ii) a translation initiation codon followed by three unique cloning sites, allowing gene fusions with a stretch of DNA coding for six histidine (iii) two translational termination codons and the ESAT-6 transcription terminator.

pIPX30 promotes high level express of β-galactosidase in mycobacteria

To characterize this novel express vector, a truncated lacZ reporter gene, generated by Asp718/PstI digestion of promoter-probe plasmid pJEM13, was inserted into the corresponding sites of pIPX30. In the resulting plasmid designated pIPX34, the lacZ gene is in frame with regard to the pIPX30 ATG initiation codon. When introduced into M. smegmatis mc2155, pIPX34 produced a high level of β-galactosidase activity (Figure 2). This level of β-galactosidase activity is comparable to the one obtained in pJN30 extracts, where lacZ is under the control of the strong pBlaF* promoter of Mycobacterium fortuitum. When transformed in Mycobacterium bovis BCG, the pIPX34 construct resulted in dark blue colonies in presence of the β-galactosidase X-gal chromogenic substrate. These observations indicated that the combination of promoter/expression cassette used in pIPX30 is functional in representative members of fast- and slow-growing mycobacteria.

Example 6: Expression and immunodetection of DES (His6) in M. smegmatis.

To validate pIPX30 as an expression/tagging vector, we expressed in this system the DES antigen of M. tuberculosis. The DES gene was recently cloned from Mycobacterium tuberculosis and encodes DES, a protein sharing conserved motifs characteristic of the class II diiron-oxoprotein family. DES is putative Δ-9 (delta 9) desaturase and could potentially be involved in the biosynthesis of mycobacterial lipids and mycolic acids. Moreover, DES is strongly recognized by sera from tuberculosis patients and represent a potential diagnostic reagent. To express DES in a mycobacterial context, the model M. smegmatis was chosen as a host because it is innocuous and can be grown to high cell density (up to 108 CFU/ml) in overnight broth cultures.

Oligonucleotides JD15 (5'-CCCGGATCCTCAGCCAAGCTGACCGACCTG-3') (SEQ ID NO: 32) and JD16 (5'-GCCGGTACCACGACGGCTCATCGCCAGTTTGCC-3') (SEQ ID NO: 33) were used to amplify by PCR the DES coding region cloned in plasmid pBS-DES. The resulting PCR fragment was digested with BAMHI and KpnI and cloned into the corresponding sites of pIPX30 to give pIPX30-DES. Protein extracts corresponding to the bacterial cell sonicate were prepared from M. smegmatis harboring pIPX30 or pIPX30-DES, and analyzed by Western blotting using anti-DES mouse polyclonal serum. A protein band migrating at about 38 kDa, was detected specifically in M. smegmatis transformed with pIPX34 plasmid but not in extracts corresponding to the pIPX30 control vector. An additional 36 kDa band detected in both protein extracts, was attributed to the endogenous M. smegmatis DES protein or alternatively to a molecule cross reacting with the anti-DES mouse serum. The same results were obtained with a commercially available monoclonal antibody directed against the (His)6 peptide, supporting the presence of six histidine at the carboxyl terminus of DES.

Example 7: Identification of LHP polypeptide in short term culture filtrate (ST-CF).

ST-CF was produced as previously described (Anderson, et al., 1991). Briefly, M. tuberculosis (8 x 106 CFU/ml) were grown in modified Sauton medium on an orbital shaker for 7 days. The culture supernatants were sterile-filtrated and concentrated on an Amicon YM3 membrane (Amicon, Danvers, MA). The ORFX protein was purified from ST-CF by preparative SDS-PAGE using the Prepcell system (BioRad, Richmond, CA). 1 ml containing 8 mg of ST-CF was applied on a matrix of 16% polyacrylamide and separation was performed under an electrical gradient for 22 hours. 3 ml fractions were collected and analyzed on silverstained SDS-PAGE. 3 ml of the fractions containing the ORFX protein was concentrated in the presence of 0.1 SDS in a Centricon-3 unit (Amicon) followed by acetone precipitation. The precipitate was redissolved in Tricine SDS-PAGE gel (Novex, San Diego, USA). After electrophoresis the gel was blotted to Problott PVDF membrane (Applied Biosystems, Foster City, CA) by semidry electroblotting in 10 mM CAPS, 10%

methanol, pH 11. The PVDF membrane was stained with 0.1% Coomassie R-250 in 40% methanol, 1% acetid acid, and destained in 50% methanol. The band of interest was excised and subjected to N-terminal sequence analysis by automated Edman degradation using a Procise 494 sequencer (Applied Biosystems) as described by the manufacturer.

By N-terminal amino acid sequencing, the inventors have obtained the following sequence. A-E-M-K-T-D-A-A-T-L-X-Q-E-A-G, wherein X represents any amino acid, said sequence corresponding to the N-terminal sequence of LHP, the methionine residue located at the NH2-terminal position having been naturally removed by the bacterial enzymatic machinery. (SEQ ID NO: 34)

hlp codes for the 10kDa culture filtrate protein CFP-10

The ESAT-6 protein consists of 95 amino acids and was previously shown to be present in the M. tuberculosis ST-CF. Since *lhp* is next to esat-6, and potentially encodes a polypeptide of 100 amino acids, we investigated its eventual presence in the M. tuberculosis ST-CT. Low molecular weight ST-CF fractions were separated by preparative SDS-PAGE and submitted to systematic N-terminal sequencing. As shown in figure 11 , fraction number 4 yielded a peptide sequence matching almost perfectly (14/15) with the N-teminus deduced from the M. *tuberculosis lhp* gene sequence. This 10 kDa culture filtrate protein was referred to as CFP-10. To further characterize the *lhp* gene product, we over-expressed and purified recombinant CFP-10 in E. coli, in fusion with a stretch of 8 histidines. Separation of rCFP-10 by SDS-PAGE indicated an apparent molecular weight of 14 kDa (Figure 12 B), slignty higher than the apparent molecular weight of recombinant ESAT-6 (His6) (10 kDa). The difference of size between native and recombinant CFP-10 may be attribuable to the presence of the histidine tag. These results demonstrated that M. *tuberculosis lhp* is a gene and encodes a small polypeptide, which like ESAT-6, is found in the low-molecular weight fraction of the ST-CF. In spite of the fact that no obvious exportation signal was identified so far in the sequence of LHP, our data suggest this protein is released

extracellularly during broth cultivation of M. tuberculosis, as already observed for ESAT-6.

Example 8 : Immunological data on CFP10 obtained from different species (Mice, Guinea pigs, cattle and humans)

Mice

Recognition during infection:

CFP10 is not very strongly recognized during infection (Approx. 1/3 of the level ST-CF).

Immunogenicity :

Immunization with CPF10/DDA induces a strong in vitro release of IFN-γ (7000 pg/ml) detected in the blood 1 week after the last immunization.

Guinea pigs

DTH-Response:

CFP10 has been tested on BCG vaccinated, M. avium and M. tub infected and naive animals. In BCG vaccinated, M. avium infected and naive animals no DTH response was measured compared to M. tub infected were a significant DTH response was observed.

Cattles

DTH-response:

CFP10 has been tested on both M. avium and M. bovis infected animals. In M. avium infected (ppdA positive) animals no DTH response was measured compared to M. bovis (ppdB positive) infected were a significant DTH response was observed in many of the cattles. Further more blood cells isolated from cattles infected with M. bovis induced an in vitro proliferative response and release of IFN-γ after stimulation with CFP10.

Human

In human only TB infected but not BCG vaccinated donors respond to CFP10.

As it appears from the teachings of the Specification, the invention is not limited in scope to one or several of the above detailed embodiments; the present invention also embraces all the alternatives that can be performed by one skilled in the same technical field, without deviating from the subject or from the scope of the instant invention.

References

1. Aguado et al., Immuno Biol., 184:113-125.

2. Andersen, P., A. B. Andersen, Sorensen, A. L. and S. Nagai. 1995. Recall of long-lived immunity to Mycobacterium tuberculosis infection in mice. J. Immunol. 154:3359-3372.

3. Andersen, P., D. Askgaard, L. Ljungqvist, W. M. Bentzon, and I. Heron. (1991) Infect. Immun. 59:1905-1910

4. Barany, F., 1911, Proc. Natl. Acad. Sci. USA, 88:189-193.

5. Bashyam, M. D. and A. K. Tyagi. 1994. An efficient and high-yielding method for isolation of RNA from mycobacteria. BioTechniques 17:834-836.

6. Bashyam, M. D., Kaushal, D., Dasgupta, K. K. and A. K. Tyagi. 1996. A study of the mycobacterial transcriptional apparatus: Identification of novel features in promoter elements. J. Bacteriol 178:4847-4853.

7. Bengard A. et al., 1994, In: Tuberculosis : Pathogenesis, Protection and Control, Barry R. Bloom, Ed., American Society for Microbiology, Washington, D.C. 20005, Chapter 21, 307-332.

8. Berthet, F.-X., Rauzier, J., Lim, E. M., Philipp, W., Giequel, B. and D. Portnoï. 1995. Characterization of the Mycobacterium tuberculosis erp gene encoding a potential cell surface protein with repetitive structures. Microbiology 141:2123-2130.

9. Bowie J. U. et al., 1990, Science, 47:1306-1310.

10. Burg J. L. et al., 1996, Mol. And Cell. Probes, 10:257-271.

11. Chu B. C. F. et al., 1986, Nucleic Acids Res., 14:5591-5603.

12. Collado-Vides, J., Magasanik, B. and J. D. Gralla. 1991. Control site location and transcriptional regulation in Escherichia coli. Microbiol. Rev. 55:371-394.

13. Collins, D. M. 1996. In search of tuberculosis virulence genes. Trends Microbiol. 4:426-430.

14. Collins, D. M., Kawakami, R. P., De Lisle, G. W., Pascopella, L., Bloom, B. R. and W. Jacobs Jr. 1995. Mutation of the principal sigma factor causes loss of

virulence in a strain of the Mycobacterium tuberculosis complex. Proc. Natl. Acad. Sci. USA 92:8036-8040.

15. Duck P. et al., 1990, Biotechniques, 9:142-147.

16. Garbe T. et al., 1993, Infect. Immun., 61(1):260-267.

17. Goldman et al., 1986, Ann. Rev. Biophys. Chem., 15:321-353.

18. Guateli J. C. et al., 1990, Proc. Natl. Acad. Sci. USA, 87:1874-1878.

19. Harboe, M., Oettinger, T., Gotten Wiker, H., Rosenkrands, I. and P. Andersen. 1996. Evidence for occurrence of the ESAT-6 protein in Mycobacterium tuberculosis and virulent Mycobacterium bovis and for its absence in Mycobacterium bovis-BCG. Infect. Immun. 64:16-22.

20. Harris D. P. et al., 1994, Infect. Immun., 62(7):2963-2972.

21. Herrmann J. L. et al., 1996, The EMBO J., 15(14):3547-3554.

22. Houbenweyl, 1974, in Meuthode der Organischen Chemie, E. Wunsch Ed., Volume 15-I et 15 II, Thieme, Stuttgart.

23. Kenney, T. J. and G. Churchward. 1996. Genetic analysis of the Mycobacterium smegmatis rpsL promoter. J. Bacteriol. 178:3564-3571.

24. Kersten G. F. et al., 1995, Biochem. Biophys. Acta, 1241:117-138.

25. Kievitis T. et al., 1991, J. Virol. Methods, 35:273-286.

26. Kremer, L., Baulard, A., Estaquier, J., Content, J., Capron, A. and C. Locht. 1995. Analysis of the Mycobacterium tuberculosis antigen 85A promoter region. J. Bacteriol. 177:642-653.

27. Kwoh D. Y. et al., 1989, Proc. Natl. Acad. Sci. USA, 86:1173-1177.

28. Kyte et al., 1982, J. Mol. Biol., 157:105-132

29. Landegren U. et al., 1988, Science, 241:1077-1080.

30. Lizardi P. M. et al., 1988, Bio Technology, 6:1197-1202.

31. Maheiras, G. G., Sabo, P. J., Hickey, M. J., Devinder, C. S. and C. K. Stover. 1996. Molecular analysis of genetic differences between Mycobacterium bovis-BCG and virulent M. bovis J. Bacteriol. 178:1274-1284.

32. Matthews J. A. et al., 1988, Anal. Biochem., 169:1-25.

33. Merrifield, 1963, J. Am. Chem. Soc., 85:2149-2154.

34. Miele E. A. et al., 1983, J. Mol. Biol., 171:281-295.

35. Morein b., 1990, Immunol. Letters, 25:281-283.

36. Mowat A. M. et al., 1991, Immunol. Today, 12:383-385.

37. Pelicic V. et al., 1996, Mol. Microbiol., 20(5):919-925.

38. Quinn, F. D., Newmann, G. W. and c. H. King. 1996. Virulence determinant of Mycobacterium tuberculosis in Tuberculosis Ed. T. Shinnik, Springer verlag.

39. Reyrat J-M. et al., 1995, Proc. Natl. Acad. Sci. USA, 92:8768-8772.

40. Rougeot, C., I. Rosinski-Chupin, e. Njamkepo, and F. Rougeon. Eur. J. Biochem. 219(3):765-773, 1994.

41. Sambrook, J. Fritsch, E. F., and T. Maniatis. 1989. Molecular cloning: a laboratory manual. 2ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

42. Sanchez-Pescador R., 1988, J. Clin. Microbiol., 26(10):1934-1938.

43. Sathish, M., Esser, R. E., Tholle, J. E. R. and J. E. Clark-Curtiss. 1990. Identification and characterization of antigenic determinants of Mycobacterium leprae that react with antibodies in sera of leprosy patients. Infect. Immun. 58:1327-1336.

44. Segev D., 1992, in << Non-radioactive Labeling and Detection of Biomolecules >>. Kessler C. Springer Verlag, Berlin, New York, 197-205.

45. Sørensen, A. L., Nagai, S., Houen, G., Andersen, P. and Å. B. Andersen. 1995. Purification and characterization of a low-molecular-mass T-cell antigen secreted by Mycobacterium tuberculosis Infect. Immun. 63:1710-1717.

46. Spargo C. A. et al., 1996, Mol. and Cell. Probes, 10:247-256

47. Stone B. B. et al., 1996, Mol. and Cell. Probes, 10:359-370.

48. Tam J. P., 1988, Proc. Natl. Acad. Sci., 85:5409-5413.

49. Timm, J., Lim, E. M. and B. Giequel. 1994. Escherichia coli-mycobacteria shuttle vectors for operon and gene fusions to lacZ: the pJEM series. J. Bacteriol. 176:6749-6753.

50. Urdea M. S. et al., 1991, Nucleic Acids Symp. Ser., 24:197-200.

51. Urdea M. S., 1988, Nucleic Acids Research, 11:4937-4957.

52. Wabiko et al., 1986, DNA, 5(4):305-314.

53. Walker G. T. et al., 1992, Nucleic Acids Res., 20:1691-1696.